KEMENY
SNELL
THOMPSON

INTRODUCTION TO FINITE MATHEMATICS

INTRODUCTION TO FINITE MATHEMATICS

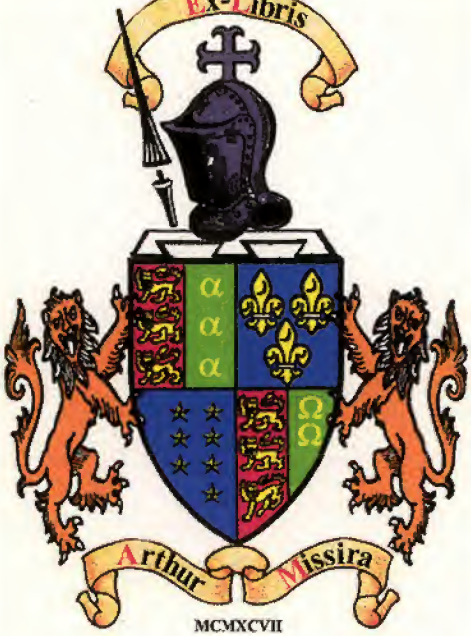
AP
KEM

PRENTICE
HALL

INTRODUCTION TO FINITE MATHEMATICS

KEMENY
SNELL
THOMPSON

56/-



One of 1600+ items free from the Internet Archive uploaded by artmisa

6567

AP
KEM.

27 SEP 1972

WITHDRAWN

This book is due for return on or before the
last date shown above.

6564

BEDGLEY PARK COLLEGE
PRESTWICH,
MANCHESTER.

INTRODUCTION TO

FINITE MATHEMATICS

PRENTICE-HALL MATHEMATICS SERIES

Dr. Albert A. Bennett, Editor

INTRODUCTION TO

FINITE MATHEMATICS

JOHN G. KEMENY

J. LAURIE SNELL

GERALD L. THOMPSON

Department of Mathematics, Dartmouth College

Englewood Cliffs, N. J., PRENTICE-HALL, INC.

© Copyright, 1956, 1957, by
PRENTICE-HALL, INC.
Englewood Cliffs, N. J.

ALL RIGHTS RESERVED. NO PART OF THIS BOOK MAY BE REPRODUCED IN ANY FORM, BY MIMEOGRAPH OR ANY OTHER MEANS, WITHOUT PERMISSION IN WRITING FROM THE PUBLISHERS.

Library of Congress Catalog Card Number: 57-7294

First printing ..... *January, 1957*
Second printing ..... *July, 1957*
Third printing ..... *April, 1958*
Fourth printing ..... *January, 1959*

PRINTED IN THE UNITED STATES OF AMERICA

48382

PREFACE

In the usual undergraduate mathematical curriculum, the courses which a student takes during his first two years are those leading up to the calculus and the calculus itself. A few years ago, the department of mathematics at Dartmouth College decided to introduce a different kind of freshman course which students could elect along with these more traditional ones. The new course was to be designed to introduce a student to some concepts in modern mathematics early in his college career. While primarily a mathematics course, it was to include applications to the biological and social sciences and thus provide a point of view, other than that given by physics, concerning the possible uses of mathematics.

In planning the proposed course, we found that there was no textbook available to fulfill our needs and, therefore, we decided to write such a book. Our aim was to choose topics which are initially close to the students' experience, which are important in modern day mathematics, and which have interesting and important applications. To guide us in the latter we asked for the opinions of a number of behavioral scientists about the kinds of mathematics a future behavioral scientist might need. The main topics of the book were chosen from this list.

Our purpose in writing the book was to develop several topics from a central point of view. In order to accomplish this on an elementary level, we restricted ourselves to the consideration of *finite* problems, that is, problems which do not involve infinite sets, limiting processes, continuity, etc. By so doing it was possible to go further into the subject matter than would otherwise be possible, and we found that the basic ideas of finite mathematics were easier to state and theorems about them considerably easier to prove than their infinite counterparts.

The first five chapters form a natural unit. The discussion of the set of logical possibilities (in Chapter I) leads to the idea of the truth set associated with a statement which, in turn, gives a natural way of defining the probability that a statement is true (in Chapter IV). The correspondence that exists among logical operations (Chapter I), set operations (Chapter II), and probability operations (Chapter IV) becomes especially transparent in the finite case. A very useful pedagogical device, that of a "tree" (a special type of diagram) is used in these chapters and in the rest of the book to illustrate and clarify ideas. In particular, this allows an introduction to the theory of stochastic processes in an elementary manner. The Markov chains here introduced help to motivate vector and matrix theory, which is presented in Chapter V.

In Chapter VI the student is introduced to two recent branches of mathematics that have proved useful in applications, namely, linear programming and the theory of games. We are able to explain the basic ideas of both relatively quickly because of the mathematical preparation given in the earlier chapters.

In our concluding Chapter VII we discuss several significant applications of mathematics to the behavioral sciences. These were selected for their interest both to mathematicians and to behavioral scientists. One topic was chosen from each of five sciences: sociology, genetics, psychology, anthropology, and economics. A reader may find it more difficult to read parts of this chapter than the earlier chapters, but it was found necessary to make it so in order that nontrivial applications could be taken up and pursued far enough to see the contributions mathematics makes. In teaching a course from our book we would not expect that all of the topics from this chapter would be used. We hope, however, that Chapter VII will serve as reference and self-study material for ambitious students.

The Committee on the undergraduate program of the Mathematical Association of America was planning a new freshman mathematics program at the same time we were planning our book. They had already written Part I of *Universal Mathematics*, which is an introduction to analytic geometry and the calculus, and were making plans for Part II. When the chairman of that committee learned of the similarity of the plans for our book to those for Part II of *Universal Mathematics*, he invited one of us to join his committee. We believe that our book agrees with the spirit of their recommendations. We

are grateful for their permission to use some of their illustrations, of which the applications to voting problems are the principal ones.

The report of the Committee on Mathematical Training of Social Scientists of the Social Science Research Council appeared after our plans were completed. We were pleased to note that on many questions we had reached the same conclusions as had that committee. They recommend two years of training, about half in the calculus and half along the lines here discussed. A semester course based on our book together with a semester of calculus would give the student a distribution in the proportions recommended by that committee.

The basic core of the book consists of the unasterisked sections of Chapters I-V. This material should be covered in every course. Flexibility is provided by the inclusion of additional material, the optional (asterisked) sections of these chapters, Chapter VI, and Chapter VII. By emphasizing the first five chapters, the course would be a basic mathematics course. By aiming at Chapter VII and taking up several of these applications, the course can be designed as a mathematics course suited for the behavioral scientist. Chapter VI is appropriate as supplementary material for either type of course. We have included a bibliography at the end of each chapter to guide those interested in further reading.

The only prerequisite for this book is the mathematical maturity obtained from two and a half or more years of high school mathematics. Our book has been tried successfully in a freshman course at Dartmouth College and for supplementary reading in other courses. It has also been used in a mathematics course for faculty members in the behavioral sciences.

We wish to thank Dartmouth College for releasing us from part of our teaching duties to enable us to prepare this book. Thanks are also due to A. W. Tucker for his valuable advice and to our colleagues in the mathematics department at Dartmouth for their many helpful suggestions. We are also grateful to James K. Schiller for reading the manuscript and for providing the reactions of a student. Finally we wish to thank Joan Snell, Margaret P. Andrews, and Stephen Russell for their invaluable aid in the preparation of the manuscript.

J. G. K.
J. L. S.
G. L. T.

CONTENTS

I. COMPOUND STATEMENTS 1

1. Purpose of the Theory 1
2. The Most Common Connectives 5
3. Other Connectives 10
*4. Statements Having Given Truth Tables 15
5. Logical Possibilities 19
6. Tree Diagrams 25
7. Logical Relations 30
*8. A Systematic Analysis of Logical Relations . . . 34
9. Variants of the Conditional 39
10. Valid Arguments 42
*11. The Indirect Method of Proof 46
*12. Applications to Switching Circuits 49

II. SETS AND SUBSETS 54

1. Introduction 54
2. Operations on Subsets 58
3. The Relationship between Sets and Compound Statements 63
*4. The Abstract Laws of Set Operations 68
5. Two-Digit Number Systems 70
*6. Voting Coalitions 74

III. PARTITIONS AND COUNTING 79

1. Partitions 79
*2. Applications 83
3. The Number of Elements in a Set 88
4. Permutations 92
5. Counting Partitions 96
6. Some Properties of the Numbers $\binom{n}{j}$ 101
7. Binomial and Multinomial Theorems 105
*8. Voting Power 108

IV. PROBABILITY THEORY 113

1. Introduction 113
2. Properties of a Probability Measure 116
3. The Equiprobable Measure 121
4. Two Nonintuitive Examples 124
5. Conditional Probability 129
*6. Measures as Areas 133
7. Finite Stochastic Processes 139
8. Independent Trials with Two Outcomes 146
*9. A Problem of Decision 151
*10. The Law of Large Numbers 156
*11. Independent Trials with More than Two Outcomes 162
12. Expected Value 166
13. Markov Chains 171

V. VECTORS AND MATRICES 178

1. Column and Row Vectors 178
2. The Product of Vectors; Examples 183
3. Matrices and Their Combination with Vectors . . 190
4. The Addition and Multiplication of Matrices . . . 198
5. The Solution of Linear Equations 205
6. The Inverse of a Square Matrix 212

7. Applications of Matrix Theory to Markov Chains . *217*
8. Examples of Markov Chains *224*
*9. Linear Functions and Transformations *233*
*10. Permutation Matrices *236*
*11. Subgroups of Permutation Groups *243*

*VI. LINEAR PROGRAMMING AND THE THEORY OF GAMES 249

1. Convex Sets *249*
2. Maxima and Minima of Linear Functions *254*
3. Linear Programming Problems *259*
4. Strictly Determined Games *266*
5. Nonstrictly Determined Games *271*
6. Matrix Games *279*
7. More on Matrix Games: The Fundamental Theorem *290*
8. $2 \times n$ and $m \times 2$ Matrix Games *293*
9. Simplified Poker *302*

*VII. APPLICATIONS TO BEHAVIORAL SCIENCE PROBLEMS 307

1. Sociometric Matrices *307*
2. Communication Networks *315*
3. Stochastic Processes in Genetics *321*
4. Absorbing Markov Chains and Genetics *326*
5. The Estes Learning Model *334*
6. Limiting Probabilities in the Estes Model *339*
7. Marriage Rules in Primitive Societies *343*
8. The Choice of Marriage Rules *348*
9. Model of an Expanding Economy *353*
10. Existence of an Economic Equilibrium *360*

INDEX 369

Chapter I

COMPOUND STATEMENTS

1. PURPOSE OF THE THEORY

A *statement* is a verbal or written assertion. In the English language such assertions are made by means of declarative sentences. For example, "It is snowing" and "I made a mistake in signing up for this course" are statements.

The two statements quoted above are *simple statements.* A combination of two or more simple statements is a *compound statement.* For example, "It is snowing, and I wish that I were out of doors, but I made the mistake of signing up for this course," is a compound statement.

It might seem natural that one should make a study of simple statements first, and then proceed to the study of compound ones. However, the reverse order has proved to be more useful. Because of the tremendous variety of simple statements, the theory of such statements is very complex. It has been found in mathematics that it is often fruitful to assume for the moment that a difficult problem has been solved and then to go on to the next problem. Therefore we shall proceed as if we knew all about simple statements and study only the way they are compounded. The latter is a relatively easy problem.

While the first systematic treatment of such problems is found in the writings of Aristotle, mathematical methods were first employed by George Boole about 100 years ago. The more polished techniques

now available are the product of twentieth century mathematical logicians.

The fundamental property of any statement is that it is either true or false (and that it cannot be both true and false). Naturally, we are interested in finding out which is the case. For a compound statement it is sufficient to know which of its components are true since the truth values (i.e., the truth or falsity) of the components determine in a way to be described later the truth value of the compound.

Our problem then is twofold: (1) In how many different ways can statements be compounded? (2) How do we determine the truth value of a compound statement given the truth values of its components?

Let us prepare our mathematical tools. In any mathematical formula we find three kinds of symbols: *constants*, *variables*, and *auxiliary symbols*. For example, in the formula $(x + y)^2$ the plus sign and the exponent are constants, the letters x and y are variables, and the parentheses are auxiliary symbols. Constants are symbols whose meanings in a given context are fixed. Thus in the formula given above, the plus sign indicates that we are to form the sum of the two numbers x and y, while the exponent 2 indicates that we are to multiply $(x + y)$ by itself. Variables always stand for entities of a given kind, but they allow us to leave open just which particular entity we have in mind. In our example above the letters x and y stand for unspecified numbers. Auxiliary symbols function somewhat like punctuation marks. Thus if we omit the parentheses in the expression above we obtain the formula $x + y^2$ which has quite a different meaning than the formula $(x + y)^2$.

In this chapter we shall use variables of only one kind. We indicate these variables by the letters p, q, r, etc., which will stand for unspecified statements. These statements frequently will be simple statements but may also be compound. In any case we know that, since each variable stands for a statement, it has an (unknown) truth value.

The constants that we shall use will stand for certain connectives used in the compounding of statements. We will have one symbol for forming the negation of a statement and several symbols for combining two statements. It will not be necessary to introduce symbols for the compounding of three or more statements, since we can show that the same combination can also be formed by compounding them two at a time. In practice only a small number of basic constants are used

and the others are defined in terms of these. It is even possible to use only a single connective! (See Section 4, Exercises 10 and 11.)

The auxiliary symbols that we shall use are, for the most part, the same ones used in elementary algebra. Any case where the usage is different will be explained.

Examples. As examples of simple statements let us take "The weather is nice" and "It is very hot." We will let p stand for the former and q for the latter.

Suppose we wish to make the compound statement that both are true "The weather is nice *and* it is very hot." We shall symbolize this statement by $p \wedge q$. The symbol $\wedge$, which can be read "and," is our first connective.

In place of the strong assertion above we might want to make the weak (cautious) assertion that one or the other of the statements is true: "The weather is nice *or* it is very hot." We symbolize this assertion by $p \vee q$. The symbol $\vee$, which can be read "or," is the second connective which we shall use.

Suppose we believed that one of the statements above was false, for example, "It is *not* very hot." Symbolically we would write $\sim q$. Our third connective is then $\sim$, which can be read "not."

More complex compound statements can now be made. For example, $p \wedge \sim q$ stands for "The weather is nice *and* it is *not* very hot."

EXERCISES

1. The following are compound statements or may be so interpreted. Find their simple components.

(a) It is hot and it is raining.

(b) It is hot but it is not very humid.

[*Ans.* "It is hot"; "it is very humid."]

(c) It is raining or it is very humid.

(d) Jack and Jill went up the hill.

(e) The murderer is Jones or Smith.

(f) It is neither necessary nor desirable.

(g) Either Jones wrote this book or Smith did not know who the author was.

2. In Exercise 1 assign letters to the various components, and write the statements in symbolic form. [*Ans.* (b) $p \wedge \sim q$.]

3. Write the following statements in symbolic form, letting p be "Fred is smart" and q be "George is smart."

(a) Fred is smart and George is stupid.
(b) George is smart and Fred is stupid.
(c) Fred and George are both stupid.
(d) Either Fred is smart or George is stupid.
(e) Neither Fred nor George is smart.
(f) Fred is not smart, but George is stupid.
(g) It is not true that Fred and George are both stupid.

4. Assume that Fred and George are both smart. Which of the seven compound statements in Exercise 3 are true?

5. Write the following statements in symbolic form.

(a) Fred likes George. (Statement p.)
(b) George likes Fred. (Statement q.)
(c) Fred and George like each other.
(d) Fred and George dislike each other.
(e) Fred likes George, but George does not reciprocate.
(f) George is liked by Fred, but Fred is disliked by George.
(g) Neither Fred nor George dislikes the other.
(h) It is not true that Fred and George dislike each other.

6. Suppose that Fred likes George and George dislikes Fred. Which of the eight statements in Exercise 5 are true?

7. For each statement in Exercise 5 give a condition under which it is false. [*Ans.* (c) Fred does not like George.]

8. Let p be "Stock prices are high," and q be "Stocks are rising." Give a verbal translation for each of the following.

(a) $p \wedge q$.
(b) $p \wedge \sim q$.
(c) $\sim p \wedge \sim q$.
(d) $p \vee \sim q$.
(e) $\sim(p \wedge q)$.
(f) $\sim(p \vee q)$.
(g) $\sim(\sim p \vee \sim q)$.

9. Using your answers to Exercise 8, parts (e), (f), (g), find simpler symbolic statements expressing the same idea.

10. Let p be "I have a dog," and q be "I have a cat." Translate into English and simplify: $\sim[\sim p \vee \sim \sim \sim q] \wedge \sim \sim \sim p$.

2. THE MOST COMMON CONNECTIVES

The truth value of a compound statement is determined by the truth values of its components. When discussing a connective we will want to know just how the truth of a compound statement made from this connective depends upon the truth of its components. A very convenient way of tabulating this dependency is by means of a *truth table.*

Let us consider the compound $p \wedge q$. Statement p could be either true or false and so could statement q. Thus there are four possible pairs of truth values for these statements and we want to know in each case whether or not the statement $p \wedge q$ is true. The answer is straightforward: If p and q are both true, then $p \wedge q$ is true, and otherwise $p \wedge q$ is false. This seems reasonable since the assertion $p \wedge q$ says no more and no less than that p and q are both true.

p	q	$p \wedge q$
T	T	T
T	F	F
F	T	F
F	F	F

Figure 1

Figure 1 gives the truth table for $p \wedge q$, the *conjunction* of p and q. The truth table contains all the information that we need to know about the connective $\wedge$, namely it tells us the truth value of the conjunction of two statements given the truth values of each of the statements.

p	q	$p \vee q$
T	T	?
T	F	T
F	T	T
F	F	F

Figure 2

We next look at the compound statement $p \vee q$, the *disjunction* of p and q. Here the assertion is that one or the other of these statements is true. Clearly, if one statement is true and the other false, then the disjunction is true, while if both statements are false, then the disjunction is certainly false. Thus we can fill in the last three rows of the truth table for disjunction (see Figure 2).

Observe that one possibility is left unsettled, namely, what happens if both components are true? Here we observe that the everyday usage of "or" is ambiguous. Does "or" mean "one or the other or both" or does it mean "one or the other but not both"?

Let us seek the answer in examples. The sentence "this summer I

will date Jean or Pat," allows for the possibility that the speaker may date both girls. However the sentence "I will go to Dartmouth or to Princeton," indicates that only one of these schools will be chosen. "I will buy a TV set or a phonograph next year," could be used in either sense; the speaker may mean that he is trying to make up his mind which one of the two to buy, but it could also mean that he will buy *at least one* of these—possibly both. We see that sometimes the context makes the meaning clear but not always.

A mathematician would never waste his time on a dispute as to which usage "should" be called the disjunction of two statements. Rather he recognizes two perfectly good usages, and calls one the *inclusive disjunction* (p or q or both) and the other the *exclusive disjunction* (p or q but not both). The symbol $\vee$ will be used for inclusive disjunction, and the symbol $\veebar$ will be used for exclusive disjunction. The truth tables for each of these are found in Figures 3 and 4 below.

p	q	$p \vee q$
T	T	T
T	F	T
F	T	T
F	F	F

Figure 3

p	q	$p \veebar q$
T	T	F
T	F	T
F	T	T
F	F	F

Figure 4

Unless we state otherwise, our disjunctions will be inclusive disjunctions.

p	$\sim p$
T	F
F	T

Figure 5

The last connective which we shall discuss in this section is *negation*. If p is a statement, the symbol $\sim p$, called the negation of p, asserts that p is false. Hence $\sim p$ is true when p is false, and false when p is true. The truth table for negation is shown in Figure 5.

Besides using these basic connectives singly to form compound statements, several can be used to form a more complicated compound statement, in much the same way that complicated algebraic expressions can be formed by means of the basic arithmetic operations. For example, $\sim(p \wedge q)$, $p \wedge \sim p$, and $(p \vee q) \vee \sim p$ are all compound statements. They

are to be read "from the inside out" in the same way that algebraic expressions are, namely, quantities inside the innermost parentheses are first grouped together, then these parentheses are grouped together, etc. Each compound statement has a truth table which can be constructed in a routine way. The following examples show how to construct truth tables.

Example 1. Consider the compound statement $p \vee \sim q$. We begin the construction of its truth table by writing in the first two columns the four possible pairs of truth values for the statements p and q. Then we write the proposition in question, leaving plenty of space between symbols so that we can fill in columns below. Next we copy the truth values of p and q in the columns below their occurrences in the proposition. This completes step 1 of the construction, see Figure 6.

p	q	$p \vee \sim q$	
T	T	T	T
T	F	T	F
F	T	F	T
F	F	F	F
Step No.		1	1

Figure 6

Next we treat the innermost compound, the negation of the variable q, completing step 2, see Figure 7.

p	q	$p \vee \sim$	q
T	T	T F	T
T	F	T T	F
F	T	F F	T
F	F	F T	F
Step No.		1 2	1

Figure 7

Finally we fill in the column under the disjunction symbol, which gives us the truth value of the compound statement for various truth values of its variables. To indicate this we place two parallel lines on each side of the final column, completing step 3 as in Figure 8.

p	q	p	$\vee$	$\sim$	q
T	T	T	T	F	T
T	F	T	T	T	F
F	T	F	F	F	T
F	F	F	T	T	F
Step No.		1	3	2	1

Figure 8

The next two examples show truth tables of more complicated compounds worked out in the same manner. There are only two basic rules which the student must remember when working these: first, work from the "inside out"; and second, the truth values of the compound statement are found in the last column filled in during this procedure.

Example 2. The truth table for the statement $(p \vee \sim q) \wedge \sim p$ together with the numbers indicating the order in which the columns are filled in appears in Figure 9.

p	q	$(p$	$\vee$	$\sim$	$q)$	$\wedge$	$\sim$	p
T	T	T	T	F	T	F	F	T
T	F	T	T	T	F	F	F	T
F	T	F	F	F	T	F	T	F
F	F	F	T	T	F	T	T	F
Step No.		1	3	2	1	4	2	1

Figure 9

Example 3. The truth table for the statement $\sim[(p \wedge q) \vee (\sim p \wedge \sim q)]$ together with the numbers indicating the order in which the columns are filled appears in Figure 10.

p	q	$\sim$	$[(p$	$\wedge$	$q)$	$\vee$	$(\sim$	p	$\wedge$	$\sim$	$q)]$
T	T	F	T	T	T	T	F	T	F	F	T
T	F	T	T	F	F	F	F	T	F	T	F
F	T	T	F	F	T	F	T	F	F	F	T
F	F	F	F	F	F	T	T	F	T	T	F
Step No.		5	1	3	1	4	2	1	3	2	1

Figure 10

EXERCISES

1. Give a compound statement which symbolically states "p or q but not both," using only $\sim$, $\vee$, and $\wedge$.

2. Construct the truth table for your answer to Exercise 1, and compare this with Figure 4.

3. Construct the truth table for the symbolic form of each statement in Exercise 3 of Section 1. How does Exercise 4 of Section 1 relate to these truth tables?

4. Construct a truth table for each of the following:

(a) $\sim(p \wedge q)$. [*Ans.* FTTT.]

(b) $p \wedge \sim p$. [*Ans.* FF.]

(c) $(p \vee q) \vee \sim p$. [*Ans.* TTTT.]

(d) $\sim[(p \vee q) \wedge (\sim p \vee \sim q)]$. [*Ans.* TFFT.]

5. Let p stand for "Jones passed the course" and q stand for "Smith passed the course" and translate into symbolic form the statement "It is not the case that Jones and Smith both failed the course." Construct a truth table for this compound statement. State *in words* the circumstances under which the statement is true.

6. Construct a simpler statement about Jones and Smith that has the same truth table as the one in Exercise 5.

7. Let $p \mid q$ express that "p and q are not both true." Write a symbolic expression for $p \mid q$ using $\sim$ and $\wedge$.

8. Write a truth table for $p \mid q$.

9. Write a truth table for $p \mid p$. [*Ans.* Same as Figure 5.]

10. Write a truth table for $(p \mid q) \mid (p \mid q)$. [*Ans.* Same as Figure 1.]

11. Construct a truth table for each of the following:

(a) $\sim(p \vee q) \vee \sim(q \vee p)$. [*Ans.* FFFT.]

(b) $\sim(p \vee q) \wedge p$. [*Ans.* FFFF.]

(c) $\sim(p \veebar q)$. [*Ans.* TFFT.]

(d) $\sim(p \mid q)$. [*Ans.* TFFF.]

12. Construct two symbolic statements, using only $\sim$, $\vee$, and $\wedge$, which have the truth tables (a) and (b), respectively:

p	q	(a)	(b)
T	T	T	T
T	F	F	F
F	T	T	F
F	F	T	T

3. OTHER CONNECTIVES

Suppose we did not wish to make an outright assertion but rather an assertion containing a condition. As examples consider the following sentences. "If the weather is nice, I will take a walk." "If the following statement is true, then I can prove the theorem." "If the cost of living continues to rise, then the government will impose rigid curbs." Each of these statements is of the form "*if* p *then* q." The *conditional* is then a new connective which is symbolized by the arrow $\rightarrow$.

Of course the precise definition of this new connective must be made by means of a truth table. If both p and q are true, then $p \rightarrow q$ is certainly true, and if p is true and q false, then $p \rightarrow q$ is certainly false. Thus the first two lines of the truth table can easily be filled in, see Figure 11a. Suppose now that p is false; how shall we fill in the

p	q	$p \rightarrow q$
T	T	T
T	F	F
F	T	?
F	F	?

Figure 11a

p	q	$p \rightarrow q$
T	T	T
T	F	F
F	T	T
F	F	T

Figure 11b

last two lines of the truth table in Figure 11a? At first thought one might suppose that it would be best to leave it completely undefined. However, to do so would violate our basic principle that a statement is either true or false.

Therefore we make the completely arbitrary decision that the conditional, $p \rightarrow q$, is *true* whenever p is false, regardless of the truth value of q. This decision enables us to complete the truth table for the conditional and it is given in Figure 11b. A glance at this truth table shows that the conditional $p \rightarrow q$ is considered false only if p is true and q is false. If we wished we might rationalize the arbitrary decision made above by saying that if statement p happens to be false then we give the conditional $p \rightarrow q$ the "benefit of the doubt" and consider it true (see Exercise 1).

In everyday conversation it is customary to combine simple statements only if they are somehow related. Thus we might say "It is raining today and I will take an umbrella," but we would not say "I read a good book and I will take an umbrella." However, the rather ill-defined concept of relatedness is difficult to enforce. Concepts related to each other in one person's mind need not be related in another's. In our study of compound statements no requirement of relatedness is imposed on two statements in order that they be compounded by any of the connectives. This freedom sometimes produces strange results in the use of the conditional. For example, according to the truth table in Figure 11b the statement "If $2 \times 2 = 5$, then black is white" is true, while the statement "If $2 \times 2 = 4$, then cows are monkeys" is false. Since we use the "if . . . then . . ." form usually only when there is a causal connection between the two statements, we might be tempted to label both of the above statements as nonsense. At this point it is important to remember that no such causal connection is intended in the usage of $\rightarrow$; the meaning of the conditional is contained in Figure 11b and nothing more is intended. This point will be discussed again in Section 7 in connection with implication.

p	q	$p \leftrightarrow q$
T	T	T
T	F	F
F	T	F
F	F	T

Figure 12

Closely connected to the conditional connective is the *biconditional* statement, $p \leftrightarrow q$, which may be read "*p if and only if q*." The bi-

conditional statement asserts that if p is true, then q is true, and if p is false, then q is false. Hence the biconditional is true in these cases and false in the others so that its truth table can be filled in as in Figure 12.

The biconditional is the last of the five connectives which we shall use in this chapter. The table below gives a summary of them together with the numbers of the figures giving their truth tables.

Name	Symbol	Translated as	Truth Table
Conjunction	$\wedge$	"and"	Figure 1
Disjunction (inclusive)	$\vee$	"or"	Figure 3
Negation	$\sim$	"not"	Figure 5
Conditional	$\rightarrow$	"if . . . then . . ."	Figure 11b
Biconditional	$\leftrightarrow$	". . . if and only if . . ."	Figure 12

Remember that the complete definition of each of these connectives is given by its truth table. The examples below show the use of the two new connectives.

Examples. In Figures 13 and 14 the truth tables of two statements are worked out following the procedure of Section 2.

p	q	p	$\rightarrow$	$(p$	$\vee$	$q)$
T	T	T	T	T	T	T
T	F	T	T	T	T	F
F	T	F	T	F	T	T
F	F	F	T	F	F	F
Step No.		1	3	1	2	1

Figure 13

It is also possible to form compound statements from three or more simple statements. The next example is a compound formed from three simple statements p, q, and r. Notice that there will be a total

p	q	$\sim$	p	$\leftrightarrow$	$(p$	$\rightarrow$	$\sim$	$q)$
T	T	F	T	T	T	F	F	T
T	F	F	T	F	T	T	T	F
F	T	T	F	T	F	T	F	T
F	F	T	F	T	F	T	T	F
Step No.		2	1	4	1	3	2	1

Figure 14

of eight possible triples of truth values for these three statements so that the truth table for our compound will have eight rows as shown in Figure 15.

p	q	r	$[p$	$\rightarrow$	$(q$	$\vee$	$r)]$	$\wedge$	$\sim$	$[p$	$\leftrightarrow$	$\sim$	$r]$
T	T	T	T	T	T	T	T	T	T	T	F	F	T
T	T	F	T	T	T	T	F	F	F	T	T	T	F
T	F	T	T	T	F	T	T	T	T	T	F	F	T
T	F	F	T	F	F	F	F	F	F	T	T	T	F
F	T	T	F	T	T	T	T	F	F	F	T	F	T
F	T	F	F	T	T	T	F	T	T	F	F	T	F
F	F	T	F	T	F	T	T	F	F	F	T	F	T
F	F	F	F	T	F	F	F	T	T	F	F	T	F
Step No.			1	3	1	2	1	5	4	1	3	2	1

Figure 15

EXERCISES

1. One way of filling in the question-marked squares in Figure 11a is given in Figure 11b. There are three other possible ways.

(a) Write the three other truth tables.

(b) Show that each one of these truth tables has an interpretation in terms of the connectives now available to us.

2. Write truth tables for $q \vee p$, $q \wedge p$, $q \rightarrow p$, $q \leftrightarrow p$. Compare these with the truth tables in Figures 3, 1, 11b, and 12, respectively.

3. Construct truth tables for:

(a) $p \rightarrow (q \vee r)$. [*Ans.* TTTFTTTT.]

(b) $(p \vee r) \wedge (p \rightarrow q)$. [*Ans.* TTFFTFTF.]

(c) $(p \vee q) \leftrightarrow (q \vee p)$. [*Ans.* TTTT.]

(d) $p \wedge \sim p$. [*Ans.* FF.]

(e) $(p \rightarrow p) \vee (p \rightarrow \sim p)$. [*Ans.* TT.]

(f) $(p \vee \sim q) \wedge r$. [*Ans.* TFTFFFTF.]

(g) $[p \rightarrow (q \rightarrow r)] \rightarrow [(p \rightarrow q) \rightarrow (p \rightarrow r)]$. [*Ans.* TTTTTTTT.]

4. For each of the following statements (i) find a symbolic form, and (ii) construct the truth table. Use the notation: p for "Joe is smart," q for "Jim is stupid," r for "Joe will get the prize."

(a) If Joe is smart and Jim is stupid, then Joe will get the prize. [*Ans.* TFTTTTTT.]

(b) Joe will get the prize if and only if either he is smart or Jim is stupid. [*Ans.* TFTFTFFT.]

(c) If Jim is stupid but Joe fails to get the prize, then Joe is not smart. [*Ans.* Same as (a).]

5. Construct truth tables for each of the following, and give an interpretation.

(a) $(p \rightarrow q) \wedge (q \rightarrow p)$. (Compare with Figure 12.)

(b) $(p \wedge q) \rightarrow p$.

(c) $q \rightarrow (p \vee q)$.

(d) $(p \rightarrow q) \leftrightarrow (\sim p \vee q)$.

6. The truth table for a statement compounded from two simple statements has four rows, and the truth table for a statement compounded from three simple statements has eight rows. How many rows would the truth table for a statement compounded from four simple statements have? How many for five? For n? Devise a systematic way of writing down these latter truth tables.

7. Let p be "It is raining," and q be "The wind is blowing." Translate each of the following into symbolic form.

(a) If it rains, then the wind blows.

(b) If the wind blows, then it rains.

(c) The wind blows if and only if it rains.

(d) If the wind blows, then it does not rain.

(e) It is not the case that the wind blows if and only if it does not rain.

8. Construct truth tables for the statements in Exercise 7. [*Ans.* TFTT, TTFT, TFFT, FTTT, TFFT.]

9. Construct a truth table for

(a) $(p \vee q) \leftrightarrow (\sim r \wedge \sim s)$.

(b) $(p \wedge q) \rightarrow \sim[\sim p \wedge (r \vee s)]$.

10. Construct a truth table for $\sim[(\sim p \wedge \sim q) \wedge (p \vee r)]$.
[*Ans.* TTTTTTFT.]

11. Find a simpler statement having the same truth table as the one found in Exercise 10.

*4. STATEMENTS HAVING GIVEN TRUTH TABLES

In the preceding two sections we showed how to construct the truth table for any compound statement. It is also interesting to consider the converse problem, namely, given a truth table to find one or more statements having this truth table. The converse problem always has a solution, and in fact, a solution using only the connectives $\wedge$, $\vee$, and $\sim$. The discussion which we give here is valid only for a truth table in three variables but can easily be extended to cover the case of n variables.

As observed in the last section a truth table with three variables has eight rows, one for each of the eight possible triples of truth values. Suppose that our given truth table has its last column consisting entirely of F's. Then it is easy to check that the truth table of the statement $p \wedge \sim p$ also has only F's in its last column, so that this statement serves as an answer to our problem. We now need consider only truth tables having one or more T's. The method that we shall use is to construct statements that are true in one case only, and then to construct the desired statement as a disjunction of these.

It is not hard to construct statements that are true in only one case. In Figure 16 are listed eight such statements, each true in exactly

p	q	r	Basic Conjunctions
T	T	T	$p \wedge q \wedge r$
T	T	F	$p \wedge q \wedge \sim r$
T	F	T	$p \wedge \sim q \wedge r$
T	F	F	$p \wedge \sim q \wedge \sim r$
F	T	T	$\sim p \wedge q \wedge r$
F	T	F	$\sim p \wedge q \wedge \sim r$
F	F	T	$\sim p \wedge \sim q \wedge r$
F	F	F	$\sim p \wedge \sim q \wedge \sim r$

Figure 16

one case. We shall call such statements *basic conjunctions*. Such a basic conjunction contains each variable or its negation, depending on whether the line on which it appears in Figure 16 has a T or an F under the variable. Observe that the disjunction of two such basic conjunctions will be true in exactly two cases, the disjunction of three in three cases, etc. Therefore, to find a statement having a given truth table simply form the disjunction of those basic conjunctions which occur in Figure 16 on the rows where the given truth table has T's.

Example 1. Find a statement whose truth table has T's in the first, second, and last rows, and F's in the other rows. The required statement is the disjunction of the first, second, and eighth basic conjunctions, that is

$$(p \wedge q \wedge r) \vee (p \wedge q \wedge \sim r) \vee (\sim p \wedge \sim q \wedge \sim r).$$

In Exercise 2 you will show that this statement has the required truth table.

Example 2. A logician is captured by a tribe of savages and placed in a jail having two exits. The savage chief offers the captive the following chance to escape: "One of the doors leads to certain death and the other to freedom. You can leave by either door. To help you in making a decision, two of my warriors will stay with you and answer any one question which you wish to ask of them. I must warn you, however, that one of my warriors is completely truthful while the other always lies." The chief then leaves, believing that he has given his captive only a sporting chance to escape.

After thinking a moment our quick-witted logician asks one question and then chooses the door leading to freedom. What question did he ask?

Let p be the statement "The first door leads to freedom," and q be the statement "You are truthful." It is clear that p and q are useless questions in themselves, so let us try compound statements. We want to ask a single question for which a "yes" answer means that p is true and a "no" answer means that p is false, regardless of which warrior is asked the question. The answers desired to these questions are listed in Figure 17.

The next thing to consider is, what would be the truth table of a question having the desired answers. If the warrior answers "yes"

p	q	Desired Answer	Truth Table of Question
T	T	yes	T
T	F	yes	F
F	T	no	F
F	F	no	T

Figure 17

and if he is truthful, that is if q is true, then the truth value is T. But if he answers "yes" and he is a liar, that is if q is false, then the truth value is F. A similar analysis holds if the answer is "no." The truth values of the desired question are shown in Figure 17.

Therefore we have reduced the problem to that of finding a statement having the truth table of Figure 17. Following the general method outlined above, we see that the statement

$$(p \wedge q) \vee (\sim p \wedge \sim q)$$

will do. Hence the logician asks the question: "Does the first door lead to freedom and are you truthful, or does the second door lead to freedom and are you lying?" The reader can show (Exercise 3) that the statement $p \leftrightarrow q$ also has the truth table given in Figure 17, hence a shorter equivalent question would be: "Does the first door lead to freedom if and only if you are truthful?"

As can be seen in Example 2, the method does not necessarily yield the simplest possible compound statement. However it has two advantages: (1) It gives us a mechanical method of finding a statement that solves the problem. (2) The statement appears in a standard form. The latter will be made use of in designing switching circuits (see Section 12).

EXERCISES

1. Show that each of the basic conjunctions in Figure 16 has a truth table consisting of one T appearing in the row in which the statement appears in Figure 16, and all the rest F's.

2. Find the truth table of the compound statement constructed in Example 1.

3. Show in Example 2 that the statement $p \leftrightarrow q$ has the truth table of Figure 17.

4. Construct one or more compound statements having each of the following truth tables, (a), (b), and (c).

p	q	r	(a)	(b)	(c)
T	T	T	T	F	T
T	T	F	F	F	T
T	F	T	T	F	T
T	F	F	F	T	F
F	T	T	F	F	T
F	T	F	F	F	T
F	F	T	T	F	F
F	F	F	F	F	T

5. Using only $\vee$, $\wedge$ and $\sim$, write a statement equivalent to each of the following:

(a) $p \leftrightarrow q$.
(b) $p \rightarrow q$.
(c) $\sim(p \rightarrow q)$.

6. Using only $\vee$ and $\sim$ write down a statement equivalent to $p \wedge q$. Use this result to prove that any truth table can be represented by means of the two connectives $\vee$ and $\sim$.

In Exercises 7-10 we will study the new connective $\downarrow$, where $p \downarrow q$ expresses "neither p nor q."

7. Construct the truth table of $p \downarrow q$.

8. Construct the truth table for $p \downarrow p$. What other compound has this truth table? [*Ans.* Same as Figure 5.]

9. Construct the truth table for $(p \downarrow q) \downarrow (p \downarrow q)$. What other compound has this truth table? [*Ans.* Same as Figure 3.]

10. Use the results of Exercises 6, 8, and 9 to show that any truth table can be represented by means of the single connective $\downarrow$.

11. Use the results of Exercises 9, 10 following Section 2 to show that any truth table can be represented by means of the single connective $|$.

12. Write down a compound of p, q, r which is true if and only if exactly one of the three components is true.

13. The "basic conjunctions" for statements having only one variable are p and $\sim p$. Discuss the various compound statements that can be formed by disjunctions of these. How do these relate to the possible truth tables for statements of one variable? What can be asserted about an arbitrary compound, no matter how long, that contains only the variable p?

[*Ans.* There are four possible truth tables.]

14. In Example 2 there is a second question, having a different truth table than that in Figure 17, which the logician can ask. What is it?

15. A student is confronted with a true-false exam, consisting of five questions. He knows that his instructor always has more true than false questions, and that he never has three questions in a row with the same answer. From the nature of the first and last questions he knows that these must have the opposite answer. The only question to which he knows the answer is number two. And this *assures* him of having all answers correct. What did he know about question two? What is the answer to the five questions? [*Ans.* TFTTF.]

5. LOGICAL POSSIBILITIES

One of the most important contributions that mathematics can make to the solution of a scientific problem is to provide an exhaustive analysis of the logical possibilities for the problem. The role of science is then to discover facts which will eliminate all but one possibility.

Given the analysis of logical possibilities, we can ask for each assertion about the problem, and for each logical possibility, whether the assertion is true in this case. Normally, for a given statement there will be many cases in which it is true and many in which it is false. Logic will be able to do no more than to point out the cases in which the statement is true. However, there are two notable exceptions, namely, a statement that is true in every logically possible case, and one that is false in every case. Here logic alone suffices to determine the truth value.

A statement that is true in every logically possible case is said to be *logically true*. The truth of such a statement follows from the meaning of the words and the form of the statement, together with the context of the problem about which the statement is made. We will see several examples of logically true statements below. A statement that is false in every logically possible case is said to be *logically false*, or to be a *self-contradiction*. For example, the conjunction of any statement

with its own negation will always be a self-contradiction, since it cannot be true under any circumstances.

Example 1. Let us consider the following problem, which is of a type often studied in probability theory. "There are two urns; the first contains two black balls and one white ball, while the second contains one black ball and two white balls; select an urn at random and draw two balls in succession from it. What is the probability that . . . ?" Without raising questions of probability, let us ask what the possibilities are. Figures 18 and 19 give us two ways of analyzing the logical possibilities.

Case	Urn	First Ball	Second Ball
1	1	black	black
2	1	black	white
3	1	white	black
4	2	black	white
5	2	white	black
6	2	white	white

Figure 18

In Figure 18 we have analyzed the possibilities as far as colors of balls drawn was concerned. Such an analysis may be sufficient for many purposes. In Figure 19 we have carried out a finer analysis, in which we distinguished between balls of the same color in an urn. For some purposes the finer analysis may be necessary. It is important to realize that the possibilities in a given problem may be analyzed in many different ways, from a very rough grouping to a highly refined one. The only requirements on an analysis of logical possibilities are: (1) That under any conceivable circumstances one and only one of these possibilities must be the case, and (2) that the analysis is fine enough so that the truth value of each statement (under consideration in the problem) is determined in each case.

It is easy to verify that both analyses (Figures 18 and 19) satisfy the first condition. Whether the rougher analysis will satisfy the

second condition depends on the nature of the problem. If we can limit ourselves to statements like "Two black balls are drawn from the first urn," then it suffices. But if we wish to consider "The first black ball is drawn after the second black ball from the second urn," then the finer analysis is needed.

Case	Urn	First Ball	Second Ball
1	1	black no. 1	black no. 2
2	1	black no. 2	black no. 1
3	1	black no. 1	white
4	1	black no. 2	white
5	1	white	black no. 1
6	1	white	black no. 2
7	2	black	white no. 1
8	2	black	white no. 2
9	2	white no. 1	black
10	2	white no. 2	black
11	2	white no. 1	white no. 2
12	2	white no. 2	white no. 1

Figure 19

Let us consider the statement "One white ball and one black are drawn." In Figure 18 this will be true in cases 2, 3, 4, and 5, false in cases 1 and 6. For any statement we will normally find a number of cases in which it is true, others in which it is false. However, an exception to this is a statement like "At most two black balls are drawn," which is true in every case, in either analysis. Hence this statement is logically true. It follows from the very definition of the problem that we cannot draw more than two balls. Hence, also, the statement "Draw three white balls," is logically false.

What the logical possibilities are for a given set of statements will depend on the context, i.e., on the problem that is being considered. Unless we know what the possibilities are, we have not understood the task before us. This does not preclude the possibility that there may be several ways of analyzing the logical possibilities. In Example 1 above, e.g., we gave two different analyses, and others could be found. In general, the question "How many cases are there in which p is true," will depend on the analysis given. (This will be of importance in our study of probability theory.) However, the logically true and logically false statements are exceptions. A statement that is logically true (false) according to one analysis will be logically true (false) according to every other analysis of the given problem.

The truth table method was used at the beginning of the chapter for analyzing logical possibilities and this is a very rough but convenient method. Suppose that we have a statement compounded from p, q, and r. There may be a hundred cases possible according to some fine analysis. But some of these can be lumped together, since it is necessary only to distinguish those cases where the truth values of the three components turn out to be different. Then we can get at most eight cases, corresponding to the eight lines of the truth table.

In the truth tables we have always assumed, tacitly, that all these cases will occur, which amounts to saying that the components are logically unrelated (see Section 8). If this assumption is satisfied, the truth table is a perfectly satisfactory means of testing for logically true (false) statements. Since the truth value of the compound depends only on the truth values of its components, no other information is relevant. Hence the eight cases suffice for the testing of logically true (false) statements.

For example, a statement of the form $p \rightarrow (p \vee q)$ will have to be true in every conceivable case. We may have a hundred cases, giving varying truth values for p and q, but every such case must correspond to one of the four truth table cases, as far as the compound is concerned. In each of these four cases the compound is true, and therefore such a statement is logically true. An example of it is "If Jones is smart, then he is smart or lucky."

However, if the components are logically related, then a truth table analysis may not be adequate. Let p be the statement "Jim is taller than Bill," while q is "Bill is taller than Jim." And consider the statement, "Either Jim is not taller than Bill or Bill is not taller than

Jim," i.e., $\sim p \vee \sim q$. If we work the truth table of this compound, we find that it is false in the first case. But this case is not logically possible, since under no circumstances can p and q both be true! Our compound is logically true, but a truth table will not show this. Had we made a careful analysis of the possibilities as to the heights of the two men, we would have found that the compound statement is true in every case. (Such relations will be considered in Section 8. This particular pair of statements will be considered in Exercise 11 in that section.)

Example 2. As a more complicated example let us consider the classification of human beings according to height, hair color, and sex that is carried out in Figure 20. Whether this analysis into 24 cases is adequate will depend on the problem. For example, if we

Case	Height	Hair Color	Sex
1	tall	blond	male
2	tall	blond	female
3	tall	brown	male
4	tall	brown	female
5	tall	black	male
6	tall	black	female
7	tall	red	male
8	tall	red	female
9	medium	blond	male
10	medium	blond	female
11	medium	brown	male
12	medium	brown	female
13	medium	black	male
14	medium	black	female
15	medium	red	male
16	medium	red	female
17	short	blond	male
18	short	blond	female
19	short	brown	male
20	short	brown	female
21	short	black	male
22	short	black	female
23	short	red	male
24	short	red	female

Figure 20

want to allow for white hair or baldness, we must have more cases.

The statement "He is a tall man," is true in cases 1, 3, 5, 7, and false in the others. "She is a woman who is neither short nor red-haired" will be true in cases 2, 4, 6, 10, 12, and 14. On the other hand, the statement "The person is tall, medium, or short," furnishes no information. It is true in every case, hence logically true. On the other hand the statement "He is a man of less than medium height, not blond, brown, or red-haired, and not a short black-haired man" is a self-contradiction.

Of all the logical possibilities, one and only one represents the facts as they are. That is, for a given person one and only one of the 24 cases is a correct description. To know which one, we need factual information. When we say that a certain statement is "true," without qualifying it, we mean that it is true in this one case. But, as we have said before, what the case actually is lies outside the domain of logic. Logic can tell us only what the circumstances (logical possibilities) are under which a statement is true.

EXERCISES

1. Prove that the negation of a logically true statement is logically false, and the negation of a logically false statement is logically true.

2. Classify the following as (i) logically true, (ii) a self-contradiction, (iii) neither.

(a) $p \leftrightarrow p$. [*Ans.* Logically true.]

(b) $p \rightarrow \sim p$.

(c) $(p \vee q) \leftrightarrow (p \wedge q)$. [*Ans.* Neither.]

(d) $(p \rightarrow \sim q) \rightarrow (q \rightarrow \sim p)$.

(e) $(p \rightarrow q) \wedge (q \rightarrow r) \wedge \sim(p \rightarrow r)$. [*Ans.* Self-contradiction.]

(f) $(p \rightarrow q) \rightarrow p$.

(g) $[(p \rightarrow q) \rightarrow p] \rightarrow p$.

3. Figure 20 gives the possible classifications of one person according to height, color of hair, and sex. How many cases do we get if we classify two people jointly? [*Ans.* 576.]

4. For each of the 24 cases in Figure 20 state whether the following statement is true: "The person has red hair, and, if the person is a woman, then she is short."

5. In Example 1, with the logical possibilities given by Figure 18, state the cases in which the following statements are true.

(a) Urn one is selected.
(b) At least one white ball is drawn.
(c) At most one white ball is drawn.
(d) If the first ball drawn is white, then the second is black.
(e) Two balls of different color are drawn if and only if urn one is selected.

6. In Example 1 give two logically true and two logically false statements (other than those in the text).

7. In a college using grades A, B, C, D, and F, how many logically possible report cards are there for a student taking four courses? [*Ans.* 625.]

8. A man has nine coins totaling 78 cents. What are the logical possibilities for the distribution of the coins? [*Hint:* There are three possibilities.]

9. In Exercise 8, which of the following statements are logically true and which are logically false?

(a) He has at least one penny. [*Ans.* Logically true.]
(b) He has at least one nickel. [*Ans.* Neither.]
(c) He has exactly two nickels. [*Ans.* Logically false.]
(d) He has exactly three nickels if and only if he has exactly one dime. [*Ans.* Logically true.]

10. In Exercise 8, we are told that the man has no nickel in his possession. What can we infer from this?

11. Two dice are rolled. Which of the following analyses satisfy the first condition for logical possibilities? What is wrong with the others?

The sum of the numbers shown is:

(a): (1) 6, (2) not 6.
(b): (1) an even number, (2) less than 6, (3) greater than 6.
(c): (1) 2, (2) 3, (3) 4, (4) more than 4.
(d): (1) 7 or 11, (2) 2, 3, or 12, (3) 4, 5, 6, 8, 9, or 10.
(e): (1) 2, 4, or 6, (2) an odd number, (3) 10 or 12.
(f): (1) less than 5 or more than 8, (2) 5 or 6, (3) 7, (4) 8.
(g): (1) more than 5 and less than 10, (2) at most 4, (3) 7, (4) 11 or 12.

[*Ans.* (a), (c), (d), (f) satisfy the condition.]

6. TREE DIAGRAMS

A very useful tool for analyzing logical possibilities is the drawing of a "tree." This device will be illustrated by several examples.

Example 1. Consider again the example in Figure 20. Suppose we let the classification proceed as follows: first consider all human

beings before classification as being all in one group; next split this large group into three subgroups by putting the short people in one group, the medium height in another, and the tall people in the third; next split up each of these subgroups into four smaller subgroups (making a total of twelve in all) according to hair color; finally, split each of these subgroups into two parts by grouping males together and females together. The final classification then divides the group of all human beings into 24 subgroups. Figure 21 gives a graphical

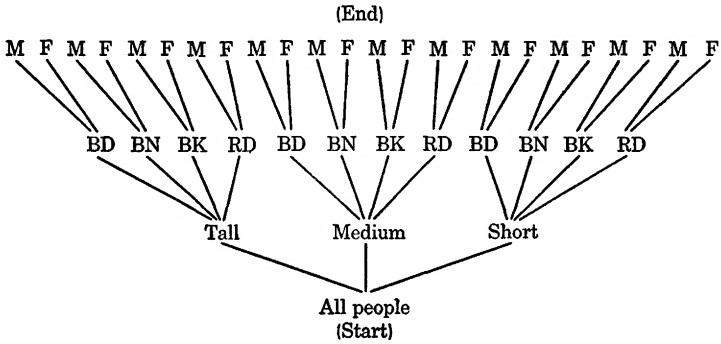


Figure 21

representation of the process described above. For obvious reasons we shall call a figure like this, which starts at a point and branches out, a *tree*.

Observe that the tree contains all the information relevant for the classification problem. Each path through the tree from the start to the end (bottom to top) represents a logical possibility. There are 24 in all, one for each end point of the tree, and similarly there are 24 cases in Figure 20. The order in which we performed the classification is arbitrary, that is, we might equally well have first classified people according to hair color, then sex, and finally height. We would still get 24 logical possibilities but the tree that we would obtain would differ from that of Figure 21 (see Exercise 1).

Example 2. Next let us consider the example of Figure 18. This is a three-stage process, first we select an urn, then draw a ball and then draw a second ball. The tree of logical possibilities is shown in

Figure 22. We note that six is the correct number of logical possibilities. The reason for this is: if we choose the first urn (which contains two black balls and one white ball) and draw from it a black ball, then the second draw may be of either color; however, if we draw a white ball first, then the second ball drawn is necessarily black. Similar remarks apply if the second urn is chosen.

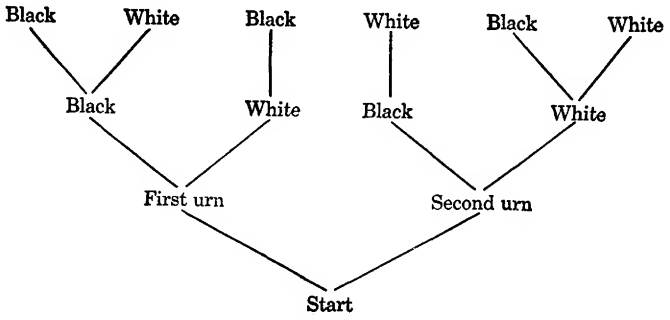


Figure 22

The student should observe that in the tree of Figure 21 each point on the same level has the same number of branches leading out of it, while in the tree of Figure 22 this is not the case.

Example 3. As a final example let us construct the tree of logical possibilities for the outcomes of a World Series played between the Dodgers and the Yankees. In Figure 23 is shown half of the tree,

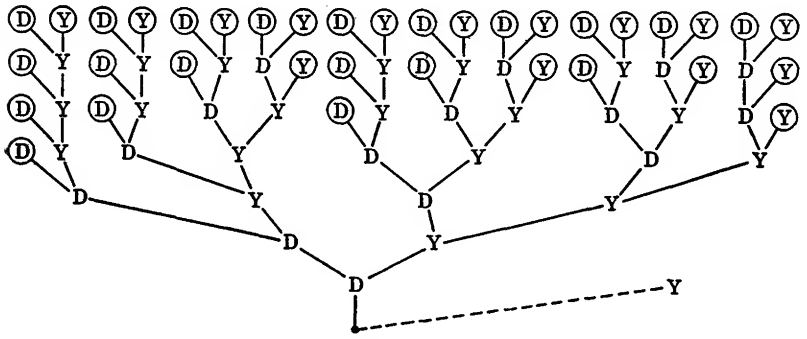


Figure 23

corresponding to the case when the Dodgers win the first game (the dotted line at the bottom leads to the other half of the tree). In the figure a 'D' stands for a Dodger win and 'Y' for a Yankee win. There are 35 possible outcomes (corresponding to the circled letters) in the half-tree shown, so that the World Series can end in 70 ways.

This example is different from the previous two in that the paths of the tree end at different levels, corresponding to the fact that the World Series ends whenever one of the teams has won four games.

Not always do we wish as detailed an analysis as that provided in the examples above. If, in Example 2, we wanted to know only the color and order in which the balls were drawn and not which urn they came from, then there would be only four logical possibilities instead of six. Then in Figure 22 the second and fourth paths (counting from the left) represent the same outcome, namely, a black ball followed by a white ball. Similarly the third and fifth paths represent the same outcome. Finally, if we cared only about the color of the balls drawn, not the order, then there are only three logical possibilities: two black balls, two white balls, or one black and one white ball.

A less detailed analysis of the possibilities for the World Series is also possible. For example we can analyze the possibilities as follows: Dodgers in 4, 5, 6, or 7 games; and Yankees in 4, 5, 6, or 7 games. The new classification reduced the number of possibilities from 70 to 8. The other possibilities have not been eliminated but merely grouped together. Thus the statement "Dodgers in 4 games" can happen in only one way, while "Dodgers in 7 games" can happen in 20 ways (see Figure 23). A still less detailed analysis would be a classification according to the number of games in the series. Here there are only four logical possibilities.

The student will find that it often requires several trials before the "best" way of listing logical possibilities is found for a given problem.

EXERCISES

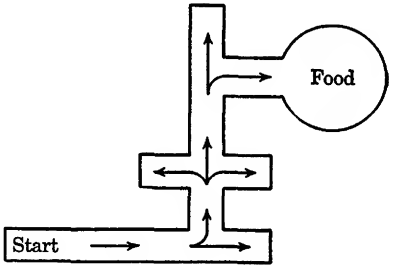
1. Construct the tree for Example 1 if the order of classification is hair color, sex, and height. Do the same if the order of classification is sex, height, and hair color. Are there any other ways of performing this classification?

2. In 1955 the Dodgers lost the first two games of the World Series, but won the series in the end. In how many ways can the series go so that the losing team wins the first two games? [*Ans.* 10.]

3. The following is a typical process in genetics: each parent has two genes for a given trait, AA or Aa or aa. The child will inherit one gene from each parent. What are the possibilities for a child if both parents are AA? What if one is AA and the other aa? What if one is AA and the other Aa? What if both are Aa? Construct a tree for each process. [Let stage one be the choice of a gene from the first parent, stage two from the second parent. Then see how many different types the resulting branches represent.]

4. It is often the case that types AA and Aa (see Exercise 3) are indistinguishable from the outside, but easily distinguishable from type aa. What are the logical possibilities if the two parents are of noticeably different types?

5. A psychologist teaches a rat to run through a maze whose shape is shown in the diagram. Let us assume that, if a rat gets into a blind alley, it goes back to the last intersection and tries a different passage than has previously ever been taken, but always in the direction of the arrows. How many possible paths are there? How many are there if a rat is stopped after making two wrong turns? [*Ans.* 20; 12.]



6. We set up an experiment similar to that of Figure 18, but urn one has two black balls and two white balls, while urn two has one white ball and four black balls. We select an urn, and draw three balls from it. Construct the tree of logical possibilities. How many cases are there? [*Ans.* 10.]

7. From the tree constructed in Exercise 6 answer the following questions.

(a) In how many cases do we draw three black balls?
(b) In how many cases do we draw two black balls and one white ball?
(c) In how many cases do we draw three white balls?
(d) How many cases does this leave? What cases are these? [*Ans.* 3.]

8. In how many ways can the World Series be played (see Figure 23) if the Dodgers win the first game and

(a) No team wins two games in a row. [*Ans.* 1.]
(b) The Dodgers win at least the odd-numbered games. [*Ans.* 5.]
(c) The winning team wins four games in a row. [*Ans.* 4.]
(d) The losing team wins four games. [*Ans.* 0.]

9. A man is considering the purchase of one of four types of stocks. Each stock may go up, go down, or stay the same after his purchase. Draw the tree of logical possibilities.

10. For the tree constructed in Exercise 9 give a statement which:

(a) Is true in half the cases.
(b) Is false in all but one case.
(c) Is true in all but one case.
(d) Is logically true.
(e) Is logically false.

11. In Exercise 6 we wish to make a rougher classification of logical possibilities. What branches (in the tree there constructed) are identified if:

(a) We do not care about the order in which the balls are drawn.
(b) We care neither about the order of balls, nor about the number of the urn selected.
(c) We care only about what urn is selected, and whether the balls drawn are all the same color.

12. Work Exercise 7 of the last section, by sketching a tree diagram.

13. A menu has a choice of soup or orange juice for an appetizer, a choice of steak, chicken, or fish for the entree, and a choice of pie or cake for dessert. A complete dinner consists of one choice in each case. Draw the tree for the possible complete dinners.

(a) How many different complete dinners are possible? [*Ans.* 12.]
(b) How many complete dinners are there which have chicken for the entree? [*Ans.* 4.]
(c) How many complete dinners are there available for a man who will eat pie only if he had steak for the entree? [*Ans.* 8.]

7. LOGICAL RELATIONS

Until now we have considered statements in isolation. Sometimes, however, we want to consider the relationship between pairs of statements. The most interesting such relation is that one statement (logically) *implies* the other one. If p implies q we also say that q follows from p, or that q is (logically) deducible from p. For example, in any mathematical theorem the hypothesis implies the conclusion.

If we have listed all logical possibilities for a pair of statements p and q, then we shall characterize implication as follows: p implies q if q is true whenever p is true, i.e., if q is true in all the logically possible cases in which p is true.

For compound statements having the same components truth ta-

bles provide a convenient method for testing this relation. In Figure 24 we illustrate this method. Let us take $p \leftrightarrow q$ as our hypothesis.

p	q	$p \leftrightarrow q$	$p \rightarrow q$	$p \vee q$
T	T	T	T	T
T	F	F	F	T
F	T	F	T	T
F	F	T	T	F

Figure 24

Since it is true only in the first and fourth cases and $p \rightarrow q$ is true in both these cases we see that the statement $p \leftrightarrow q$ implies $p \rightarrow q$. On the other hand the statement $p \vee q$ is false in the fourth case and hence it is not implied by $p \leftrightarrow q$. Again, a comparison of the last two columns of Figure 24 shows that the statement $p \rightarrow q$ does not imply and is not implied by $p \vee q$.

The relation of implication has a close affinity to the conditional statement, but it is important not to confuse the two. The conditional is a new *statement* compounded from two given statements, while implication is a *relation* between the two statements. The connection is the following: p implies q if and only if the conditional $p \rightarrow q$ is logically true.

That this is the case is shown by a simple argument. The statement p implies the statement q if q is true whenever p is true. This means that there is no case in which p is true and q false, i.e., no case in which $p \rightarrow q$ is false. But this in turn means that $p \rightarrow q$ is logically true. In Exercise 1 this result will be applied to Figure 24.

Let us now take up the "paradoxes" of the conditional. Conditional statements sound paradoxical when the components are not related. For example, it sounds strange to say that "If it is a nice day then chalk is made of wood," is true on a rainy day. It must be remembered that the conditional statement just quoted means no more and no less than that one of the following holds: (1) It is a nice day and chalk is made of wood, or (2) It is not a nice day and chalk is made of wood, or (3) It is not a nice day and chalk is not made of wood. [See Figure 11b.] And on a rainy day number 3 happens to be correct.

But it is by no means true that "It is a nice day," implies that

"Chalk is made of wood." It is logically possible for the former to be true and for the latter to be false (indeed, this is the case on a nice day, with the usual method of chalk manufacture), hence the implication does not hold. Thus while the conditional quoted in the previous paragraph is true on a given day, it is not logically true.

In common parlance "if . . . then . . ." is usually asserted on logical grounds. Hence any usage in which such an assertion happens to be true, but is not logically true, sounds paradoxical. Similar remarks apply to the common usage of "if and only if."

If the biconditional $p \leftrightarrow q$ is not only true but logically true, then this establishes a relation between p and q. Since $p \leftrightarrow q$ is true in every logically possible case, the statements p and q have the same truth value in every case. We say, under these circumstances, that p and q are (logically) *equivalent.* For compound statements having the same components, the truth table provides a convenient means of testing for equivalence. We merely have to verify that the compounds have the same truth table. Figure 25 establishes that $p \rightarrow q$ is equivalent to $\sim p \vee q$.

p	q	$p \rightarrow q$	$\sim p \vee q$
T	T	T	T
T	F	F	F
F	T	T	T
F	F	T	T

Figure 25

EXERCISES

1. Show that $(p \leftrightarrow q) \rightarrow (p \rightarrow q)$ is logically true, but that $(p \leftrightarrow q) \rightarrow (p \vee q)$ is not logically true.

2. Prove that p is equivalent to q just in case p implies q and q implies p.

3. Construct truth tables for the following compounds, and test for implications and equivalences.

(a) $p \wedge q$.
(b) $p \rightarrow \sim q$.
(c) $\sim p \vee \sim q$.
(d) $\sim p \vee q$.
(e) $p \wedge \sim q$. [*Ans.* (b) equiv. (c); (a) impl. (d); (e) impl. (b), (c).]

4. Construct truth tables for the following compounds, and arrange them in order so that each compound implies all the following ones.

(a) $\sim p \leftrightarrow q$.
(b) $p \rightarrow (\sim p \rightarrow q)$.
(c) $\sim[p \rightarrow (q \rightarrow p)]$.
(d) $p \vee q$.
(e) $\sim p \wedge q$. [*Ans.* (c), (e), (a), (d), (b).]

5. Construct a compound equivalent to $p \wedge q$, using only the connectives $\sim$ and $\vee$.

6. Construct a compound equivalent to $p \leftrightarrow q$ using only the connectives $\rightarrow$ and $\wedge$. (Cf. Exercise 2.)

7. Construct a compound statement equivalent to $p \vee q$, using only the connectives $\sim$ and $\wedge$.

8. If p is logically true, prove that:
(a) $p \vee q$ is logically true.
(b) $\sim p \wedge q$ is logically false.
(c) $p \wedge q$ is equivalent to q.
(d) $\sim p \vee q$ is equivalent to q.

9. If p and q are logically true and r is logically false, what is the status of $(p \vee \sim q) \wedge \sim r$? [*Ans.* Logically true.]

10. Prove that the conjunction or disjunction of a statement with itself is equivalent to the statement.

11. Prove that the double negation of a statement is equivalent to the statement.

12. Prove that a statement which implies its own negation is a self-contradiction.

13. What is the status of a statement equivalent to its own negation?

14. What relation exists between two logically true statements? Between two self-contradictions?

15. Prove that a logically true statement is implied by every statement, and that a self-contradiction implies every statement.

16. Using the results of Section 4, Exercises 10 and 11, prove that for any compound statement there is an equivalent compound statement:
(a) Whose only connective is $\downarrow$.
(b) Whose only connective is $|$.

*8. A SYSTEMATIC ANALYSIS OF LOGICAL RELATIONS

The relation of implication is characterized by the fact that it is impossible for the hypothesis to be true and the conclusion to be false. If two statements are equivalent, it is impossible for one to be true and the other to be false. Thus we see that for an implication one truth table case must not occur, and for an equivalence two of the four truth table cases must not occur. The absence of one or more truth table cases is thus characteristic of logical relations. In this section we shall investigate all conceivable relations that can exist between two statements.

We shall say that two statements are *unrelated* if each of the four truth table cases (see Figure 26) can occur. The two statements are *related* if one or more of the four cases in Figure 26 cannot occur. [Cf. Section 5.]

p	q	Case No.
T	T	1
T	F	2
F	T	3
F	F	4

Figure 26

If p and q are statements such that exactly one of the cases in Figure 26 is excluded, then we say that there is a *onefold* relation between them. Obviously there are four possible onefold relations which we list below. (a) If case 1 is excluded, the two statements cannot both be true. In this case p and q are said to be a pair of *contraries* or are said to be *inconsistent.* (b) If case 2 is excluded, then (cf. Section 7) p implies q. (c) If case 3 is excluded, it is false that q is true and p is false, that is, q implies p. (d) If case 4 is excluded, both statements cannot be false, i.e., one of them is true. Such a pair of statements is called a pair of *subcontraries.*

If p and q are statements such that exactly two of the cases in Figure 26 are excluded, then we say that there is a *twofold* relation between them. There are six ways in which two cases can be selected from four, but several of these do not produce interesting relations. For example, suppose cases 1 and 2 are excluded; then p cannot be true, i.e., it is logically false. Similarly, if cases 1 and 3 are excluded, then q is logically false. On the other hand, if cases 3 and 4 are excluded, then p is logically true; and if 2 and 4 are excluded, then q is logically true. Hence we see that these choices do not give us new

relations; they merely indicate that one of the two statements is logically true or false. We now have only two alternatives remaining: (A) cases 2 and 3 are excluded which means that the two statements are equivalent; and (B) cases 1 and 4 are excluded, which means that the two statements cannot both be true and cannot both be false; in other words, one must be true and the other false. We shall then say that p and q are *contradictories*.

It is not hard to see that there are no threefold relations, for if three of the cases in Figure 26 are excluded, then there is only one possibility for each of the two statements, so that each must be either logically true or logically false.

We have already discussed implication and equivalence and have noted their connection to the conditional and the biconditional, respectively. We can do the same for the three remaining relations. If p and q are subcontraries, then they cannot both be false; since this is the only case in which their disjunction is false, we see that p and q are subcontraries if and only if $p \vee q$ is logically true. If p and q are contraries, then they cannot both be true; since this is the only case in which their conjunction is true, we see that p and q are contraries if and only if $p \wedge q$ is logically false. Finally, if p and q are contradictories, then cases 1 and 4 of Figure 26 are excluded, hence $p \leftrightarrow q$ is logically false. (Note also that, if p and q are contradictories, then $p \veebar q$ is logically true.) The table in Figure 27 gives a summary of the relevant facts about the six relations we have derived.

Case(s) Excluded	Relation	Alternate Definition
T-T	Contraries	$p \wedge q$ logically false
F-F	Subcontraries	$p \vee q$ logically true
T-F	First implies second	$p \rightarrow q$ logically true
F-T	Second implies first	$q \rightarrow p$ logically true
T-F and F-T	Equivalents	$p \leftrightarrow q$ logically true
T-T and F-F	Contradictories	$p \leftrightarrow q$ logically false

Figure 27

Subcontraries are not of great theoretical importance, but contraries and contradictories are very important. Each of these relations can be generalized to hold for more than two statements. If we

have n different statements, not all of which can be true, then we say that they are *inconsistent.* Then the conjunction of these statements must be false. Special cases of inconsistent statements are the following: if $n = 1$, then we have a single self-contradictory statement; and if $n = 2$, then we have a pair of inconsistent statements (i.e., a pair of contraries).

If we have n different statements such that one and only one of them can be true, then we say they form a *complete set of alternatives.* Again the special cases are: if $n = 1$, then we have a single logically true statement; and if $n = 2$, then we have a pair of contradictories.

Truth tables again furnish a method for recognizing when relations hold between statements. The examples below show how the method works.

Examples. Consider the five compound statements, all having the same components, which appear in Figure 28. Find all relations which exist between pairs of these statements.

p	q	$p \wedge q$	$\sim p \vee \sim q$	$\sim p \vee q$	$\sim p$	$p \rightarrow q$
T	T	T	F	T	F	T
T	F	F	T	F	F	F
F	T	F	T	T	T	T
F	F	F	T	T	T	T
Statement Number		1	2	3	4	5

Figure 28

First of all we note that statements 3 and 5 have identical truth tables, hence they are equivalent. Therefore we need consider only one of them, say statement 3. Statements 1 and 2 have exactly opposite truth tables, hence they are contradictories. Upon comparing statements 1 and 3 we find no T-F case, so that 1 implies 3. Since numbers 1 and 4 are never both true they are contraries, while numbers 2 and 3 are never both false, so that they are subcontraries. Finally, upon comparing either 2 or 3 to 4 we find no F-T case and hence both are implied by 4. Thus the six relations we found above

are all exemplified in Figure 28. Observe also that statements p and q give an example of a pair of unrelated statements. [Cf. Section 5.]

EXERCISES

1. Construct truth tables for the following four statements and state what relation (if any) holds between each of the six pairs formable.

(a) $\sim p$.
(b) $\sim q$.
(c) $p \wedge \sim q$.
(d) $\sim(\sim p \vee q)$.

[*Ans.* (a) and (b) unrelated; (a) and (c), (d) contraries; (c), (d) imply (b); (c) equiv. (d).]

2. Construct truth tables for each of the following six statements. Give an example of an unrelated pair, and an example of each of the six possible relations among these.

(a) $p \leftrightarrow q$.
(b) $p \rightarrow q$.
(c) $\sim p \wedge \sim q$.
(d) $(p \wedge q) \vee (\sim p \wedge \sim q)$.
(e) $\sim q$.
(f) $p \wedge \sim q$.

3. Prove the following assertions:

(a) The disjunction of two contradictory statements is logically true.
(b) Two statements are equivalent if and only if either one implies the other one.
(c) The contradictories of two contraries are subcontraries.

4. What is the relation between the following pair of statements?

(a) $p \rightarrow [p \wedge \sim(q \vee r)]$.
(b) $\sim p \vee (\sim q \wedge \sim r)$. [*Ans.* Equivalent.]

5. At most how many of the following assertions can one person consistently believe?

(a) Joe is smart.
(b) Joe is unlucky.
(c) Joe is lucky but not smart.
(d) If Joe is smart, then he is unlucky.
(e) Joe is smart if and only if he is lucky.
(f) Either Joe is smart, or he is lucky, but not both. [*Ans.* 4.]

6. Prove the following assertions.

(a) The contradictories of two equivalent statements are equivalent.

(b) In a complete set of alternatives any two statements are contraries.
(c) If p and q are subcontraries, and if each implies r, then r is logically true.

7. Pick out a complete set of (four) alternatives from:
(a) It is raining but the wind is not blowing.
(b) It rains if and only if the wind blows.
(c) It is not the case that it rains and the wind blows.
(d) It is raining and the wind is blowing.
(e) It is neither raining nor is the wind blowing.
(f) It is not the case that it is raining or the wind is not blowing.
[*Ans.* (a), (d), (e), (f).]

8. What is the relation between $[p \vee \sim(q \vee r) \vee (p \wedge s)]$ and $\sim(p \wedge q \wedge r \wedge s)$? [*Ans.* Subcontraries.]

9. Suppose that p and q are contraries. What is the relation between
(a) p and $\sim q$.
(b) $\sim p$ and q.
(c) $\sim p$ and $\sim q$.
(d) p and $\sim p$.

10. Let p, q, and r be three statements such that any two of them are unrelated. Discuss the possible relations among the three statements. [*Hint:* If we ignore the order of the statements, there are 14 such relations. The relations are at most fourfold. There are two fourfold relations, and the other relations are found from these by excluding fewer cases.]

11. In Section 5 we considered an example comparing the height of two men. Suppose that we allow for the possibilities: below 5 ft 9 in., 5 ft 9 in., 5 ft 10 in., 5 ft 11 in., 6 ft 0 in., above 6 ft. We will, for the purpose of this problem, consider two men of the same height if they fall into the same category according to the above analysis.
(a) Construct the set of all possibilities for a pair of men, Jim and Bill.
(b) Find the cases in which "Jim is taller than Bill" is true.
(c) Find the cases where "Bill is taller than Jim" is true.
(d) Are all four truth table cases present?
(e) What is the relation between the two statements?

12. Construct the set of logical possibilities which classify a person with respect to sex and marital status.
(a) Show that "if the person is a bachelor then he is unmarried" is logically true.
(b) Show that "if a person is an old maid then the person is a man" is not logically false.

(c) Find the relation between "the person is a man" and "the person is a bachelor."

(d) Find a simple statement that is a subcontrary of "the person is a man," and is consistent with it.

9. VARIANTS OF THE CONDITIONAL

The conditional of two statements differs from the biconditional and from disjunctions and conjunctions of these two in that it lacks symmetry. Thus $p \vee q$ is equivalent to $q \vee p$, $p \wedge q$ is equivalent to $q \wedge p$, and $p \leftrightarrow q$ is equivalent to $q \leftrightarrow p$; *but* $p \rightarrow q$ is *not equivalent to* $q \rightarrow p$. The latter statement, $q \rightarrow p$, is called the *converse* of $p \rightarrow q$. Many of the most common fallacies in thinking arise from a confusion of a statement with its converse.

It is also of interest to consider conditionals formed from the statements p and q. The truth tables of these four conditionals together with their names are tabulated in Figure 29. We note that $p \rightarrow q$ is

		Conditional	Converse of Conditional	Converse of Contrapositive	Contrapositive
p	q	$p \rightarrow q$	$q \rightarrow p$	$\sim p \rightarrow \sim q$	$\sim q \rightarrow \sim p$
T	T	T	T	T	T
T	F	F	T	T	F
F	T	T	F	F	T
F	F	T	T	T	T

Figure 29

equivalent to $\sim q \rightarrow \sim p$. The latter is called the *contrapositive* of the former. For many arguments the contrapositive is a very useful form of the conditional. In the same manner the statement $\sim p \rightarrow \sim q$ is the converse of the contrapositive. Since the contrapositive is equivalent to $p \rightarrow q$, the converse of the former is equivalent to the converse of the latter as can be seen in Figure 29.

The use of conditionals seems to cause more trouble than the use of the other connectives, perhaps because of the lack of symmetry,

but also perhaps because there are so many different ways of expressing conditionals. In many cases only a careful analysis of a conditional statement shows whether the person making the assertion means the given conditional or its converse. Indeed, sometimes he means both of these, i.e., he means the biconditional. (See Exercise 5.)

The statement, "I will go for a walk only if the sun shines," is a variant of a conditional statement. A statement of the form "p only if q" is closely related to the statement "If p then q," but just how? Actually the two express the same idea. The statement "p only if q" states that "If $\sim q$ then $\sim p$" and hence is equivalent to "If p then q." Thus the statement at the beginning of the paragraph is equivalent to the statement, "If I go for a walk, then the sun will be shining."

Other phrases, in common use by mathematicians, which indicate a conditional statement are, "a necessary condition" and "a sufficient condition." To say that p is a sufficient condition for q means that if p takes place, then q will also take place. Hence the sentence "p is a sufficient condition for q" is equivalent to the sentence "If p then q."

Similarly, the sentence "p is a necessary condition for q" is equivalent to "q only if p." Since we know that the latter is equivalent to "If q then p," it follows that the assertion of a necessary condition is the converse of the assertion of a sufficient condition.

Finally, if both a conditional statement and its converse are asserted, then effectively the biconditional statement is being asserted. Hence the assertion "p is a necessary and sufficient condition for q" is equivalent to the assertion "p if and only if q."

EXERCISES

1. Let p stand for "I will pass this course" and q for "I will do homework regularly." Put the following statements into symbolic form.

(a) I will pass the course only if I do homework regularly.
(b) Doing homework regularly is a necessary condition for me to pass this course.
(c) Passing this course is a sufficient condition for me to do homework regularly.
(d) I will pass this course if and only if I do homework regularly.
(e) Doing homework regularly is a necessary and sufficient condition for me to pass this course.

2. Take the statement in part (a) of the previous exercise. Form its

converse, its contrapositive, and the converse of the contrapositive. For each of these give both a verbal and a symbolic form.

3. Let p stand for "It snows," and q for "The train is late." Put the following statements into symbolic form.

(a) Snowing is a sufficient condition for the train to be late.
(b) Snowing is a necessary and sufficient condition for the train to be late.
(c) The train is late only if it snows.

4. Take the statement in part (a) of the previous exercise. Form its converse, its contrapositive, and the converse of its contrapositive. Give a verbal form of each of them.

5. Prove that the conjunction of a conditional and its converse is equivalent to the biconditional.

6. To what is the conjunction of the contrapositive and its converse equivalent? Prove it.

7. Prove that

(a) $\sim\sim p$ is equivalent to p.
(b) The contrapositive of the contrapositive is equivalent to the original conditional.

8. "For a matrix to have an inverse it is necessary that its determinant be different from zero." Which of the following statements follow from this? [No knowledge of matrices is required.]

(a) For a matrix to have an inverse it is sufficient that its determinant be zero.
(b) For its determinant to be different from zero it is sufficient for the matrix to have an inverse.
(c) For its determinant to be zero it is necessary that the matrix have no inverse.
(d) A matrix has an inverse if and only if its determinant is not zero.
(e) A matrix has a zero determinant only if it has no inverse.

[*Ans.* (b), (c), (e).]

9. "A function that is differentiable is continuous." This statement is true for all functions, but its converse is not always true. Which of the following statements are true for all functions? [No knowledge of functions is required.]

(a) A function is differentiable only if it is continuous.
(b) A function is continuous only if it is differentiable.
(c) Being differentiable is a necessary condition for a function to be continuous.

(d) Being differentiable is a sufficient condition for a function to be continuous.

(e) Being differentiable is a necessary and sufficient condition for a function to be differentiable. [*Ans.* (a), (d), (e).]

10. Prove that the negation of, "p is a necessary and sufficient condition for q," is equivalent to, "p is a necessary and sufficient condition for $\sim q$."

10. VALID ARGUMENTS

One of the most important tasks of a logician is the checking of *arguments*. By an argument we shall mean the assertion that a certain statement (the *conclusion*) follows from other statements (the *premises*). An argument will be said to be *valid* if and only if the conjunction of the premises implies the conclusion, i.e., whenever the premises are all true, the conclusion is also true.

It is important to realize that the truth of the conclusion is irrelevant as far as the test of the validity of the argument goes. A true conclusion is neither necessary nor sufficient for the validity of the argument. The two examples below show this, and they also show the form in which we shall state arguments, i.e., first we state the premises, then draw a line and then state the conclusion.

Example 1.

If the United States is a democracy, then its citizens have the right to vote.
Its citizens do have the right to vote.

Therefore the United States is a democracy.

The conclusion is, of course, true. However, the argument is not valid since the conclusion does not follow from the two premises.

Example 2.

In a democracy the chief executive is elected directly by the people.
In England the Prime Minister is the chief executive.
The British Prime Minister is not directly elected.

Therefore England is not a democracy.

Here the conclusion is false, but the argument is valid since the conclusion follows from the premises. If we observe that the first premise is false, the paradox disappears. There is nothing surprising

in the correct derivation of a false conclusion from false premises.

If an argument is valid, then the conjunction of the premises implies the conclusion. Hence if all the premises are true, then the conclusion is also true. However, if one or more of the premises is false, so that the conjunction of all the premises is false, then the conclusion may be either true or false. In fact all the premises could be false, the conclusion true, and the argument valid, as the following example shows.

Example 3.

All dogs have two legs.
All two-legged animals are carnivorous.
Therefore, all dogs are carnivorous.

Here the argument is valid and the conclusion is true, but both premises are false!

Each of these examples underlines the fact that neither the truth value nor the content of the statements appearing in an argument affect the validity of the argument. In Figures 30 and 31 are two valid forms of arguments:

$p \to q$
p
$\therefore q$

Figure 30

$p \to q$
$\sim q$
$\therefore \sim p$

Figure 31

The symbol $\therefore$ means "therefore." The truth tables for these argument forms appear in Figure 32.

p	q	$p \to q$	p	q	$p \to q$	$\sim q$	$\sim p$
T	T	T	T	T	T	F	F
T	F	F	T	F	F	T	F
F	T	T	F	T	T	F	T
F	F	T	F	F	T	T	T

Figure 32

For the argument of Figure 30, we see in Figure 32 that there is only one case in which both premises are true, namely, the first case, and in this case the conclusion is true, hence the argument is valid. Similarly, in the argument of Figure 31, both premises are true in the

fourth case only, and in this case the conclusion is also true, hence the argument is valid.

An argument that is not valid is called a *fallacy*. Two examples of fallacies are the following argument forms:

Fallacies

$$\begin{array}{l} p \to q \\ \underline{q} \\ \therefore p \end{array} \qquad\qquad \begin{array}{l} p \to q \\ \underline{\sim p} \\ \therefore \sim q \end{array}$$

In the first fallacy both premises are true in the first and third cases of Figure 32, but the conclusion is false in the third case, so that the argument is invalid. (This is the form of Example 1.) Similarly, in the second fallacy we see that both premises are true in the last two cases, but the conclusion is false in the third case.

We say that an argument depends only upon its form in that it does not matter what the components of the argument are. The truth tables in Figure 32 show that if both premises are true, then the conclusions of the arguments in Figures 30 and 31 are also true. For the fallacies above, the truth tables show that it is possible to choose both premises true without making the conclusion true, namely, choose a false p and a true q.

Example 4. Consider the following argument:

$$\begin{array}{l} p \to q \\ \underline{q \to r} \\ \therefore p \to r \end{array}$$

The truth table of the argument appears in Figure 33.

p	q	r	$p \to q$	$q \to r$	$p \to r$
T	T	T	T	T	T
T	T	F	T	F	F
T	F	T	F	T	T
T	F	F	F	T	F
F	T	T	T	T	T
F	T	F	T	F	T
F	F	T	T	T	T
F	F	F	T	T	T

Figure 33

Both premises are true in the first, fifth, seventh, and eighth rows of the truth table. Since in each of these cases the conclusion is also true, the argument is valid. (Example 3 can be written in this form.)

EXERCISES

1. Test the validity of the following arguments:

(a) $p \leftrightarrow q$, p, $\therefore q$ (b) $p \vee q$, $\sim p$, $\therefore q$ (c) $p \wedge q$, $\sim p \rightarrow q$, $\therefore \sim q$

[*Ans.* (a), (b) are valid.]

2. Test the validity of the following arguments:

(a) $p \rightarrow q$, $\sim q \rightarrow \sim r$, $\therefore r \rightarrow p$ (b) $p \rightarrow q$, $\sim r \rightarrow \sim q$, $\therefore \sim r \rightarrow \sim p$

[*Ans.* (b) is valid.]

3. Test the validity of the argument

$$\begin{array}{l} p \leftrightarrow q \\ q \vee r \\ \underline{\sim r} \\ \therefore \sim p \end{array}$$

[*Ans.* Not valid.]

4. Test the validity of the argument

$$\begin{array}{l} p \vee q \\ \sim q \rightarrow r \\ \underline{\sim p \vee \sim r} \\ \therefore \sim p \end{array}$$

5. Test the validity of the argument

$$\begin{array}{l} p \rightarrow q \\ \sim p \rightarrow \sim q \\ \underline{p \wedge \sim r} \\ \therefore s \end{array}$$

6. Given are the premises $\sim p \rightarrow q$ and $\sim r \rightarrow \sim q$. We wish to find a valid conclusion involving p and r (if there is any).

(a) Construct truth tables for the two premises.

(b) Note the cases in which the conclusion must be true.

(c) Construct a truth table for a combination of p and r only, filling in T wherever necessary.

(d) Fill in the remainder of the truth table, making sure that you do not end up with a logically true statement.

(e) What combination of p and r has this truth table? This is a valid conclusion. [*Ans.* $p \vee r$.]

7. Translate the following argument into symbolic form, and test its validity.

If this is a good course, then it is worth taking.
Either the grading is lenient, or the course is not worth taking.
But the grading is not lenient.
Therefore, this is not a good course. [*Ans.* Valid.]

8. Write the following argument in symbolic form, and test its validity:

"For the candidate to win it is sufficient that he carry New York. He will carry New York only if he takes a strong stand on civil rights. He will not take a strong stand on civil rights. Therefore, he will not win."

9. Write the following argument in symbolic form and test its validity:

"Father praises me only if I can be proud of myself. Either I do well in sports or I cannot be proud of myself. If I study hard, then I cannot do well in sports. Therefore, if father praises me, then I do not study hard."

10. Supply a conclusion to the following argument, making it a valid argument. [Adapted from Lewis Carroll.]

"If he goes to a party, he does not fail to brush his hair.
To look fascinating it is necessary to be tidy.
If he is an opium eater, then he has no self-command.
If he brushes his hair, he looks fascinating.
He wears white kid gloves only if he goes to a party.
Having no self-command is sufficient to make one look untidy.
Therefore . . ."

*11. THE INDIRECT METHOD OF PROOF

A proof is an argument which shows that a conditional statement of the form $p \rightarrow q$ is logically true. (Namely, p is the conjunction of the premises, and q is the conclusion of the argument.) Sometimes it is more convenient to show that an equivalent conditional statement is logically true.

Example 1. Let x and y be positive integers.

Theorem. If xy is an odd number, then x and y are both odd.

Proof. Suppose, on the contrary, that they are not both odd. Then one of them is even, say $x = 2z$. Then $xy = 2zy$ is an even number, contrary to hypothesis. Hence we have proved our theorem.

Example 2. "He did not know the first name of the president of the Jones Corporation, hence he cannot be an employee of that firm. Why? Because every employee of that firm calls the boss by his first name (behind his back). Therefore, if he were really an employee of Jones, then he would know Jones's first name."

These are simple examples of a very common form of argument, frequently used both in mathematics and in everyday discussions. Let us try to unravel the form of the argument.

Given:	xy is an odd number.	He doesn't know Jones's first name.	p
To prove:	x and y are both odd numbers.	He doesn't work for Jones.	q
Suppose:	x and y are not both odd numbers.	He does work for Jones.	$\sim q$
Then:	xy is an even number.	He must know what Jones's first name is.	$\sim p$

In each case we assume the contradictory to the conclusion and derive, by a valid argument, a result contradictory to the hypothesis. This is one form of the *indirect* method of proof.

To restate, what we want to do is to show that the conditional $p \rightarrow q$ is logically true; what we actually show is that the contrapositive $\sim q \rightarrow \sim p$ is logically true. Since these two statements are equivalent our procedure is valid.

There are several other important variants of this method of proof. It is easy to check that the following statements have the same truth table as (that is, are equivalent to) the conditional $p \rightarrow q$:

(1) $(p \wedge \sim q) \rightarrow \sim p$,
(2) $(p \wedge \sim q) \rightarrow q$,
(3) $(p \wedge \sim q) \rightarrow (r \wedge \sim r)$.

The first of these shows that in the indirect method of proof we may

make use of the original hypothesis in addition to the contradictory assumption $\sim q$. The second shows that we may also use this double hypothesis in the direct proof of the conclusion q. The third shows that if, from the double hypothesis p and $\sim q$ we can arrive at a contradiction of the form $r \wedge \sim r$, then the proof of the original statement is complete. This last form of the method is often referred to as *reductio ad absurdum*.

These last forms of the method are very useful for the following reasons: First of all we see that we can always take $\sim q$ as a hypothesis in addition to p. Secondly we see that besides q there are two other conclusions ($\sim p$ or a contradiction) which are just as good.

EXERCISES

1. Construct indirect proofs for the following assertions:
(a) If x^2 is odd, then x is odd (x an integer).
(b) If I am to pass this course, I must do homework regularly.
(c) If he earns a great deal of money (more than \$20,000), he is not a college professor.

2. Give a symbolic analysis of the following argument:

"If he is to succeed, he must be both competent and lucky. Because, if he is not competent, then it is impossible for him to succeed. If he is not lucky, something is sure to go wrong."

3. Construct indirect proofs for the following assertions:
(a) If $p \vee q$ and $\sim q$, then p.
(b) If $p \leftrightarrow q$ and $q \rightarrow \sim r$ and r, then $\sim p$.

4. Give a symbolic analysis of the following argument:

"If Jones is the murderer, then he knows the exact time of death and the murder weapon. Therefore, if he does not know the exact time or does not know the weapon, then he is not the murderer."

5. Verify that forms (1), (2), and (3) given above are equivalent to $p \rightarrow q$.

6. Give an example of an indirect proof of some statement in which from p and $\sim q$ a contradiction is derived.

7. Give a statement equivalent to $(p \wedge q) \rightarrow r$, which is in terms of $\sim p$, $\sim q$, and $\sim r$. Show how this can be used in a proof where there are two hypotheses given.

*12. APPLICATIONS TO SWITCHING CIRCUITS

The theory of compound statements has many applications to subjects other than pure mathematics. As an example we shall develop a theory of simple switching networks.

A switching network is an arrangement of wires and switches which connect together two terminals T_1 and T_2. Each switch can be either "open" or "closed." An open switch prevents the flow of current, while a closed switch permits flow. The problem that we want to solve is the following: given a network and given the knowledge of which switches are closed, determine whether or not current will flow from T_1 to T_2.

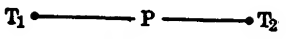


Figure 34

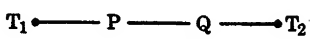


Figure 35

Figure 34 shows the simplest kind of a network in which the terminals are connected by a single wire containing a switch P. If P is closed, then current will flow between the terminals, and otherwise it does not. The network in Figure 35 has two switches P and Q in "series." Here the current flows only if both P and Q are closed.

To see how our logical analysis can be used to solve the problem stated above let us associate a statement with each switch. Let p be the statement "Switch P is closed" and let q be the statement "Switch Q is closed." Then in Figure 34 current will flow if and only if p is true. Similarly in Figure 35 the current will flow if and only if both p and q are true, that is, if and only if $p \wedge q$ is true. Thus the first circuit is represented by p and the second by $p \wedge q$.

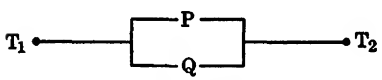


Figure 36

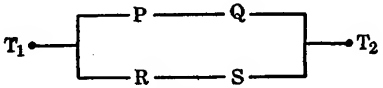


Figure 37

In Figure 36 is shown a network with switches P and Q in "parallel." In this case the current flows if either of the switches is closed, so the circuit is represented by the statement $p \vee q$.

The network in Figure 37 combines the series and parallel types

of connections. The upper branch of the network is represented by the statement $p \wedge q$ and the lower by $r \wedge s$; hence the entire circuit is represented by $(p \wedge q) \vee (r \wedge s)$. Since there are four switches and each one can be either open or closed, there are $2^4 = 16$ possible settings for these switches. Similarly, the statement $(p \wedge q) \vee (r \wedge s)$ has four variables, so that its truth table has 16 rows in it. The switch settings for which current flows correspond to the entries in the truth table for which the above compound statement is true.

Switches need not always act independently of each other. It is possible to couple two or more switches together so that they open and close simultaneously, and we shall indicate this in diagrams by giving all such switches the same letter. It is also possible to couple two switches together, so that if one is closed, the other is open. We shall indicate this by giving the first switch the letter P and the second the letter P′. Then the statement "P is closed" is true if and only if the statement "P′ is closed" is false. Therefore if p is the statement "P is closed," then $\sim p$ is the statement "P′ is closed."

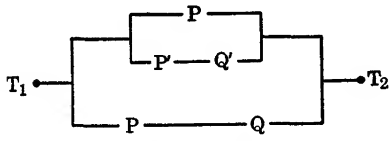


Figure 38

Such a circuit is illustrated in Figure 38. The associated compound statement is $[p \vee (\sim p \wedge \sim q)] \vee [p \wedge q]$. Since this statement is false only if p is false and q is true, the current will flow unless P is open and Q is closed. We can also check directly. If P is closed, current will flow through the top branch regardless of Q's setting. If both switches are open, then P′ and Q′ will be closed, so that current will flow through the middle branch. But if P is open and Q is closed, none of the branches will pass current.

Notice that we never had to consider current flow through the bottom branch. The logical counterpart of this fact is that the statement associated with the network is equivalent to $[p \vee (\sim p \wedge \sim q)]$ whose associated network is just the upper two branches of Figure 38. Thus the electrical properties of the circuit of Figure 38 would be the same if the lower branch were omitted.

As a last problem we shall consider the design of a switching network having certain specified properties. An equivalent problem, which we solved in Section 4, is that of constructing a compound statement having a given truth table. As in that section, we shall

limit ourselves to statements having three variables, although our methods could easily be extended.

In Section 4 we developed a general method for finding a statement having a given truth table not consisting entirely of F's. (The circuit which corresponds to a statement whose truth table consists entirely of F's is one in which current never flows, and hence is not of interest.) Each such statement could be constructed as a disjunction of basic conjunctions. Since the basic conjunctions were of the form $p \wedge q \wedge r$, $p \wedge q \wedge \sim r$, etc., each will be represented by a circuit consisting of three switches in series and will be called a *basic series circuit.* The disjunction of certain of these basic conjunctions will then be represented by the circuit obtained by putting several basic series circuits in parallel. The resulting network will not, in general, be the simplest possible such network fulfilling the requirements, but the method always suffices to find one.

Example. A three-man committee wishes to employ an electric circuit to record a secret simple majority vote. Design a circuit so that each member can push a button for his "yes" vote (not push it for a "no" vote), and so that a signal light will go on if a majority of the committee members vote yes.

p	q	r	Desired Truth Value	Corresponding Basic Conjunction
T	T	T	T	$p \wedge q \wedge r$
T	T	F	T	$p \wedge q \wedge \sim r$
T	F	T	T	$p \wedge \sim q \wedge r$
T	F	F	F	$p \wedge \sim q \wedge \sim r$
F	T	T	T	$\sim p \wedge q \wedge r$
F	T	F	F	$\sim p \wedge q \wedge \sim r$
F	F	T	F	$\sim p \wedge \sim q \wedge r$
F	F	F	F	$\sim p \wedge \sim q \wedge \sim r$

Figure 39

Let p be the statement "committee member 1 votes yes," let q be the statement "member 2 votes yes," and let r be "member 3 votes yes." The truth table of the statement "majority of the members

vote yes" appears in Figure 39. From that figure we can read off the desired compound statement as

$$(p \wedge q \wedge r) \vee (p \wedge q \wedge \sim r) \vee (p \wedge \sim q \wedge r) \vee (\sim p \wedge q \wedge r).$$

The circuit desired for the voting procedure appears in Figure 40.

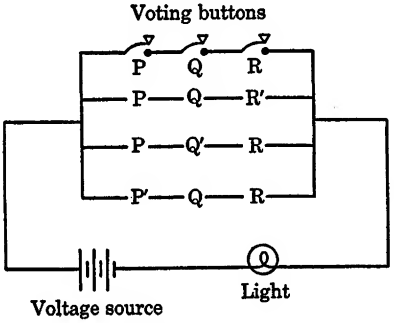


Figure 40

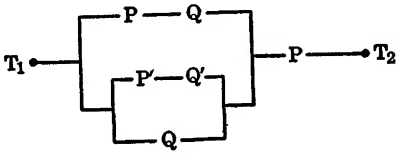
EXERCISES

1. What kind of a circuit has a logically true statement assigned to it? Give an example.

2. Construct a network corresponding to

$$[(p \wedge \sim q) \vee (\sim p \wedge q)] \vee (\sim p \wedge \sim q).$$

3. What compound statement represents:



4. Work out the truth table of the statement in Exercise 3. What does this tell us about the circuit?

5. Design a simpler circuit than the one in Exercise 3, having the same properties.

6. Construct a network corresponding to

$$[(p \vee q) \wedge \sim r] \vee [(\sim p \wedge r) \vee q].$$

7. Design a circuit for an electrical version of the game of matching pennies: At a given signal each of the two players either opens or closes a switch under his control. If they both do the same, A wins; if they do the opposite, then B wins. Design the circuit so that a light goes on if A wins.

8. In a large hall it is desired to turn the lights on or off from any one of four switches on the four walls. This can be accomplished by designing a circuit which turns the light on if an even number of switches are closed, and off if an odd number are closed. (Why does this solve the problem?) Design such a circuit.

9. A committee has five members. It takes a majority vote to carry a measure, except that the chairman has a veto (i.e., the measure carries only if he votes for it). Design a circuit for the committee, so that each member votes for a measure by pressing a button, and the light goes on if and only if the measure is carried.

10. A group of candidates is asked to take a true-false exam, with four questions. Design a circuit such that a candidate can push the buttons of those questions to which he wants to answer "true," and that the circuit will indicate the number of correct answers. [*Hint:* Have five lights, corresponding to 0, 1, 2, 3, 4 correct answers, respectively.]

11. Devise a scheme for working truth tables by means of switching circuits.

SUGGESTED READING

Tarski, A., *Introduction to Logic,* Oxford University Press, New York, 2d rev. ed., 1946, Chapters I, II.

Cohen, M. R., and E. Nagel, *An Introduction to Logic and Scientific Method,* Harcourt, Brace, New York, 1934.

Church, A., *Introduction to Mathematical Logic,* Volume I, Princeton University Press, 1956.

Hilbert, D., and W. Ackermann, *Principles of Mathematical Logic,* Chelsea, New York, 2d ed., 1950.

Allendoerfer, C. B., and C. O. Oakley, *Principles of Mathematics,* McGraw-Hill, New York, 1955, Chapter I.

Johnstone, H. W., Jr., *Elementary Deductive Logic,* Crowell, New York, 1954. Parts One and Two.

Hohn, Franz, "Some Mathematical Aspects of Switching," *The American Mathematical Monthly,* **62** (1955), pp. 75-90.

Chapter II

SETS AND SUBSETS

1. INTRODUCTION

A well-defined collection of objects is known as a *set*. This concept, in its complete generality, is of great importance in mathematics since all of mathematics can be developed by starting from it.

The various pieces of furniture in a given room form a set. So do the books in a given library, or the integers between 1 and 1,000,000, or all the ideas that mankind has had, or the human beings alive between one billion B.C. and ten billion A.D. These examples are all examples of *finite* sets, that is, sets having a finite number of elements. All the sets discussed in this book will be finite sets.

There are two essentially different ways of specifying a set. One can give a rule by which it can be determined whether or not a given object is a member of the set, or one can give a complete list of the elements in the set. We shall say that the former is a *description* of the set and the latter is a *listing* of the set. For example, we can define a set of four people as (a) the members of the quartet which played in town last night, or (b) the people whose names are Jones, Smith, Brown, and Green. It is customary to use braces to surround the listing of a set; thus the set above should be listed {Jones, Smith, Brown, Green}.

We shall frequently be interested in sets of logical possibilities, since the analysis of such sets is very often a major task in the solving of a problem. Suppose, for example, that we were interested in the successes of three candidates who enter the presidential primaries (we

assume there are no other entries). Suppose that the key primaries will be held in New Hampshire, Minnesota, Wisconsin, and California. Assume that candidate A enters all the primaries, that B does not contest in New Hampshire's primary, and C does not contest in Wisconsin's. A list of the logical possibilities is given in Figure 1. Since the New Hampshire and Wisconsin primaries can each end in two ways, and the Minnesota and California primaries can each end in three ways, there are in all $2 \cdot 2 \cdot 3 \cdot 3 = 36$ different logical possibilities as listed in Figure 1.

A set that consists of some members of another set is called a *subset* of that set. For example, the set of those logical possibilities in Figure 1 for which the statement "Candidate A wins at least three primaries" is true, is a subset of the set of all logical possibilities. This subset can also be defined by listing its members: {P1, P2, P3, P4, P7, P13, P19}.

In order to discuss all the subsets of a given set, let us introduce the following terminology. We shall call the original set the *universal set*, one-element subsets will be called *unit sets*, and the set which contains no members the *empty set*. We do not introduce special names for other kinds of subsets of the universal set. As an example, let the universal set $\mathfrak{U}$ consist of the three elements $\{a, b, c\}$. The *proper subsets* of $\mathfrak{U}$ are those sets containing some but not all of the elements of $\mathfrak{U}$. The proper subsets consist of three two-element sets, namely, $\{a, b\}$, $\{a, c\}$, and $\{b, c\}$ and three unit sets, namely, $\{a\}$, $\{b\}$, and $\{c\}$. To complete the picture we also consider the universal set a subset (but not a proper subset) of itself, and we consider the empty set $\mathfrak{E}$, that contains no elements of $\mathfrak{U}$, as a subset of $\mathfrak{U}$. At first it may seem strange that we should include the sets $\mathfrak{U}$ and $\mathfrak{E}$ as subsets of $\mathfrak{U}$, but the reasons for their inclusion will become clear later.

We saw that the three element set above had $8 = 2^3$ subsets. In general, a set with n elements has 2^n subsets, as can be seen in the following manner. We form subsets P of $\mathfrak{U}$ by considering each of the elements of $\mathfrak{U}$ in turn and deciding whether or not to include it in the subset P. If we decide to put every element of $\mathfrak{U}$ into P we get the universal set, and if we decide to put no element of $\mathfrak{U}$ into P we get the empty set. In most cases we will put some but not all the elements into P and thus obtain a proper subset of $\mathfrak{U}$. We have to make n decisions, one for each element of the set, and for each decision we have to choose between two alternatives. We can make these decisions in $2 \cdot 2 \cdot \ldots \cdot 2 = 2^n$ ways, and hence this is the number of

Possibility Number	Winner in New Hampshire	Winner in Minnesota	Winner in Wisconsin	Winner in California
P1	A	A	A	A
P2	A	A	A	B
P3	A	A	A	C
P4	A	A	B	A
P5	A	A	B	B
P6	A	A	B	C
P7	A	B	A	A
P8	A	B	A	B
P9	A	B	A	C
P10	A	B	B	A
P11	A	B	B	B
P12	A	B	B	C
P13	A	C	A	A
P14	A	C	A	B
P15	A	C	A	C
P16	A	C	B	A
P17	A	C	B	B
P18	A	C	B	C
P19	C	A	A	A
P20	C	A	A	B
P21	C	A	A	C
P22	C	A	B	A
P23	C	A	B	B
P24	C	A	B	C
P25	C	B	A	A
P26	C	B	A	B
P27	C	B	A	C
P28	C	B	B	A
P29	C	B	B	B
P30	C	B	B	C
P31	C	C	A	A
P32	C	C	A	B
P33	C	C	A	C
P34	C	C	B	A
P35	C	C	B	B
P36	C	C	B	C

Figure 1

different subsets of $\mathfrak{U}$ that can be formed. Observe that our formula would not have been so simple if we had not included the universal set and the empty set as subsets of $\mathfrak{U}$.

In the example of the voting primaries above there are 2^{36} or about 70 billion subsets. Of course, we cannot deal with this many subsets in a practical problem, but fortunately we are usually interested in only a few of the subsets. The most interesting subsets are those which can be defined by means of a simple rule such as "The set of all logical possibilities in which C loses at least two primaries." It would be difficult to give a simple description for the subset containing the elements {P1, P4, P14, P30, P34}. On the other hand, we shall see in the next section how to define new subsets in terms of subsets already defined.

Examples. We illustrate the two different ways of specifying sets in terms of the primary voting example. Let the universal set $\mathfrak{U}$ be the logical possibilities given in Figure 1.

1. What is the subset of $\mathfrak{U}$ in which candidate B wins more primaries than either of the other candidates? *Answer:* {P11, P12, P17, P23, P26, P28, P29}.

2. What is the subset in which the primaries are split two and two? *Answer:* {P5, P8, P10, P15, P21, P30, P31, P35}.

3. Describe the set {P1, P4, P19, P22}. *Answer:* The set of possibilities for which A wins in Minnesota and California.

4. How can we describe the set {P18, P24, P27}? *Answer:* The set of possibilities for which C wins in California, and the other primaries are split three ways.

EXERCISES

1. In the primary example, give a listing for each of the following sets.
   (a) The set in which C wins at least two primaries.
   (b) The set in which the first three primaries are won by the same candidate.
   (c) The set in which B wins all four primaries.

2. The primaries are considered decisive if a candidate can win three primaries, or if he wins two primaries including California. List the set in which the primaries are decisive.

3. Give simple descriptions for the following sets (referring to the primary example).

(a) {P33, P36}.
(b) {P10, P11, P12, P28, P29, P30}.
(c) {P6, P20, P22}.

4. Joe, Jim, Pete, Mary, and Peg are to be photographed. They want to line up so that boys and girls alternate. List the set of all possibilities.

5. In Exercise 4, list the following subsets.

(a) The set in which Pete and Mary are next to each other.
(b) The set in which Peg is between Joe and Jim.
(c) The set in which Jim is in the middle.
(d) The set in which Mary is in the middle.
(e) The set in which a boy is at each end.

6. Pick out all pairs in Exercise 5 in which one set is a subset of the other.

7. A TV producer is planning a half-hour show. He wants to have a combination of comedy, music, and commercials. If each is allotted a multiple of five minutes, construct the set of possible distributions of time. (Consider only the total time allotted to each.)

8. In Exercise 7, list the following subsets.

(a) The set in which more time is devoted to comedy than to music.
(b) The set in which no more time is devoted to commercials than to either music or comedy.
(c) The set in which exactly five minutes is devoted to music.
(d) The set in which all three of the above conditions are satisfied.

9. In Exercise 8, find two sets, each of which is a proper subset of the set in (a) and also of the set in (c).

2. OPERATIONS ON SUBSETS

In Chapter I we considered the ways in which one could form new statements from given statements. Now we shall consider an analogous procedure, the formation of new sets from given sets. We shall assume that each of the sets that we use in the combination is a subset of some universal set, and we shall also want the newly formed set to be a subset of the same universal set. As usual we can specify a newly formed set either by a description or by a listing.

If P and Q are two sets we shall define a new set $P \cap Q$, called the *intersection* of P and Q as follows: $P \cap Q$ is the set which contains those and only those elements which belong to both P and Q. As an

example, consider the logical possibilities listed in Figure 1. Let P be the subset in which candidate A wins at least three primaries, i.e., the set {P1, P2, P3, P4, P7, P13, P19}; let Q be the subset in which A wins the first two primaries, i.e., the set {P1, P2, P3, P4, P5, P6}. Then the intersection $P \cap Q$ is the set in which both events take place, i.e., where A wins the first two primaries *and* wins at least three primaries. Thus $P \cap Q$ is the set {P1, P2, P3, P4}.

If P and Q are two sets we shall define a new set $P \cup Q$ called the *union* of P and Q as follows: $P \cup Q$ is the set that contains those and only those elements that belong either to P or to Q (or to both). In the example in the paragraph above, the union $P \cup Q$ is the set of possibilities for which either A wins the first two primaries *or* wins at least three primaries, i.e., the set {P1, P2, P3, P4, P5, P6, P7, P13, P19}.

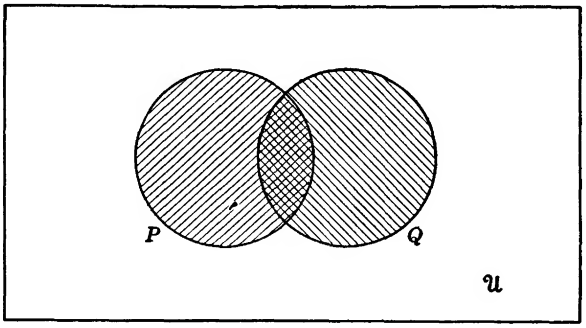


Figure 2

To help in visualizing these operations we shall draw diagrams, called Venn diagrams, which illustrate them. We let the universal set be a rectangle and let subsets be circles drawn inside the rectangle. In Figure 2 we show two sets P and Q as shaded circles. Then the doubly crosshatched area is the intersection $P \cap Q$ and the total shaded area is the union $P \cup Q$.

If P is a given subset of the universal set $\mathfrak{U}$, we can define a new set $\tilde{P}$ called the *complement* of P as follows: $\tilde{P}$ is the set of all elements of $\mathfrak{U}$ that are *not* contained in P. For example, if, as above, Q is the set in which candidate A wins the first two primaries, then $\tilde{Q}$ is the set {P7, P8, . . . , P36}. The shaded area in Figure 3 is the complement of the set P. Observe that the complement of the empty set ε

is the universal set $\mathfrak{U}$, and also that the complement of the universal set is the empty set.

Sometimes we shall be interested in only part of the complement of a set. For example, we might wish to consider the part of the complement of the set Q that is contained in P, i.e., the set $P \cap \tilde{Q}$. The shaded area in Figure 4 is $P \cap \tilde{Q}$.

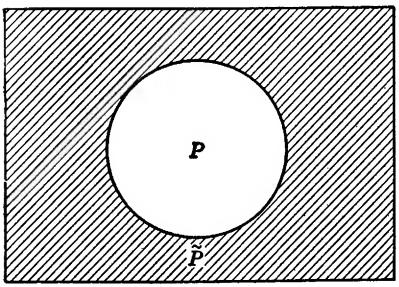


Figure 3

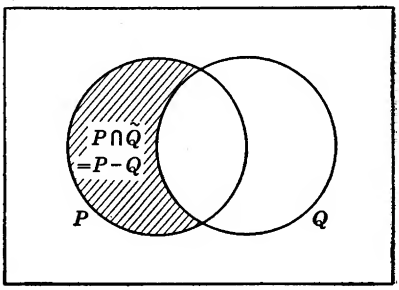


Figure 4

A somewhat more suggestive definition of this set can be given as follows: let $P - Q$ be the *difference* of P and Q, that is, the set that contains those elements of P that do not belong to Q. Figure 4 shows that $P \cap \tilde{Q}$ and $P - Q$ are the same set. In the primary voting example above the set $P - Q$ can be listed as {P7, P13, P19}.

The complement of a subset is a special case of a difference set, since we can write $\tilde{Q} = \mathfrak{U} - Q$. If P and Q are nonempty subsets whose intersection is the empty set, i.e., $P \cap Q = \mathcal{E}$, then we say that they are *disjoint* subsets.

Examples. In the primary voting example let R be the set in which A wins the first three primaries, i.e., the set {P1, P2, P3}; let S be the set in which A wins the last two primaries, i.e., the set {P1, P7, P13, P19, P25, P31}. Then $R \cap S$ = {P1} is the set in which A wins the first three primaries and also the last two, that is he wins all the primaries. We also have

$$R \cup S = \{\text{P1, P2, P3, P7, P13, P19, P25, P31}\},$$

which can be described as the set in which A wins the first three primaries or the last two. The set in which A does not win the first three primaries is $\tilde{R}$ = {P4, P5, . . . , P36}. Finally, we see that the

difference set $R - S$ is the set in which A wins the first three primaries but not both of the last two. This set can be found by taking from R the element {P1} which it has in common with S, so that $R - S = \{P2, P3\}$.

EXERCISES

1. Draw Venn diagrams for $P \cap Q$, $P \cap \tilde{Q}$, $\tilde{P} \cap Q$, $\tilde{P} \cap \tilde{Q}$.

2. Give a step-by-step construction of the diagram for $(\tilde{P} - Q) \cup (P \cap \tilde{Q})$.

3. Venn diagrams are also useful when three subsets are given. Construct such a diagram, given the subsets P, Q, and R. Identify each of the eight resulting areas in terms of P, Q, and R.

4. In testing blood, three types of antigens are looked for: A, B, and Rh. Every person is classified doubly. He is Rh positive if he has the Rh antigen, and Rh negative otherwise. He is type AB, A, or B depending on which of the other antigens he has, with type O having neither A nor B. Draw a Venn diagram, and identify each of the eight areas.

5. Considering only two subsets, the set X of people having antigen A, and the set Y of people having antigen B, define (symbolically) the types AB, A, B, and O.

6. A person can receive blood from another person if he has all the antigens of the donor. Describe in terms of X and Y the sets of people who can give to each of the four types. Identify these sets in terms of blood types.

7.

	Liked very much	Liked slightly	Disliked slightly	Disliked very much
Men	1	3	5	10
Women	6	8	3	1
Boys	5	5	3	2
Girls	8	5	1	1

This tabulation records the reaction of a number of spectators to a television show. All the categories can be defined in terms of the following four:

M (male), G (grown-up), L (liked), Vm (very much). How many people fall into each of the following categories:

(a) M. [*Ans.* 34.]
(b) $\tilde{L}$.
(c) Vm.
(d) $M \cap \tilde{G} \cap \tilde{L} \cap Vm$. [*Ans.* 2.]
(e) $\tilde{M} \cap G \cap L$.
(f) $(M \cap G) \cup (L \cap Vm)$.
(g) $\widetilde{(M \cap G)}$. [*Ans.* 48.]
(h) $(\tilde{M} \cup \tilde{G})$.
(i) $(M - G)$.
(j) $[\tilde{M} - (G \cap L \cap \widetilde{Vm})]$.

8. In a survey of 100 students, the numbers studying various languages were found to be: Spanish, 28; German, 30; French, 42; Spanish and German, 8; Spanish and French, 10; German and French, 5; all three languages, 3.

(a) How many students were studying no language? [*Ans.* 20.]
(b) How many students had French as their only language? [*Ans.* 30.]
(c) How many students studied German if and only if they studied French? [*Ans.* 38.]

[*Hint:* Draw a Venn diagram with three circles, for French, German, and Spanish students. Fill in the numbers in each of the eight areas, using the data given above. Start from the end of the list and work back.]

9. In a later survey of the 100 students (see Exercise 8), numbers studying the various languages were found to be: German only, 18; German but not Spanish, 23; German and French, 8; German, 26; French, 48; French and Spanish, 8; no language, 24.

(a) How many students took Spanish? [*Ans.* 18.]
(b) How many took German and Spanish but not French? [*Ans.* None.]
(c) How many took French if and only if they did not take Spanish? [*Ans.* 50.]

10. The report of one survey of the 100 students (see Exercise 8) stated that the numbers studying the various languages were: all three languages, 5; German and Spanish, 10; French and Spanish, 8; German and French, 20; Spanish, 30; German, 23; French, 50. The surveyor who turned in this report was fired. Why?

11. A recent survey of 100 Dartmouth students has revealed the information about their dates that is summarized in the following table.

	Beautiful and Intelligent	Plain and Intelligent	Beautiful and Dumb	Plain and Dumb
Blonde	6	9	10	20
Brunette	7	11	15	9
Redhead	2	3	8	0

Let BL = blondes, BR = brunettes, R = redheads, BE = beautiful girls, D = dumb girls. Determine the number of girls in each of the following classes.

(a) $BL \cap BE \cap D$. [*Ans.* 10.]

(b) BR.

(c) $R \cap \tilde{D}$.

(d) $(BR \cup R) \cap (BE \cup \tilde{D})$. [*Ans.* 46.]

(e) $\widetilde{BL} \cup (\widetilde{BE} \cap D)$.

12. In Exercise 11, which set of each of the following pairs has more girls as members?

(a) $(\widetilde{BL \cup BR})$ or R.

(b) $D \cap \widetilde{BE}$ or $BL - (D \cap \widetilde{BE})$.

(c) $\mathcal{E}$ or $R \cap \widetilde{BE} \cap D$.

3. THE RELATIONSHIP BETWEEN SETS AND COMPOUND STATEMENTS

The reader may have observed several times in the preceding sections that there was a close connection between sets and statements, and between set operations and compounding operations. In this section we shall formalize these relationships.

If we have a number of statements under consideration there is a natural way of assigning a set to each one of these statements. First we form the set of all logical possibilities for the statements under consideration and call this set the universal set. Then to each statement we assign the subset of logical possibilities of the universal set for which that statement is true. This idea is so important that we embody it in a formal definition.

DEFINITION. Let $p, q, r, \ldots$ be statements and let $\mathfrak{U}$ be their set of logical possibilities; let $P, Q, R, \ldots$ be the subsets of $\mathfrak{U}$ for which statements $p, q, r, \ldots$ are respectively true; then we call $P, Q, R, \ldots$ the *truth sets* of statements $p, q, r, \ldots$.

If p and q are statements, then $p \vee q$ and $p \wedge q$ are also statements and hence must have truth sets. To find the truth set of $p \vee q$ we observe that it is true whenever p is true or q is true (or both). Therefore we must assign to $p \vee q$ the logical possibilities which are in P or in Q (or both); that is, we must assign to $p \vee q$ the set $P \cup Q$. On the other hand, the statement $p \wedge q$ is true only when both p and q are true, so that we must assign to $p \wedge q$ the set $P \cap Q$.

Thus we see that there is a close connection between the logical operation of disjunction and the set operation of union, and also between conjunction and intersection. A careful examination of the definitions of union and intersection shows that the word "or" occurs in the definition of union and the word "and" occurs in the definition of intersection. Thus the connection between the two theories is not surprising.

Since the connective "not" occurs in the definition of the complement of a set, it is not surprising that the truth set of $\sim p$ is $\tilde{P}$. This follows since $\sim p$ is true when p is false, so that the truth set of $\sim p$ contains all logical possibilities for which p is false, that is, the truth set of $\sim p$ is $\tilde{P}$.

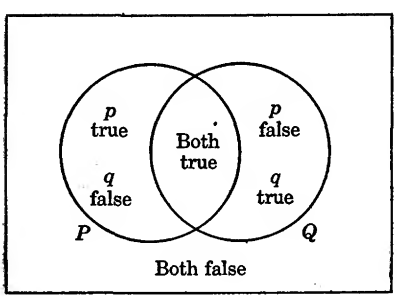


Figure 5

The truth sets of two propositions p and q are shown in Figure 5. Also marked on the diagram are the various logical possibilities for these two statements. The reader should pick out in this diagram the truth sets of the statements $p \vee q$, $p \wedge q$, $\sim p$, and $\sim q$.

The connection between a statement and its truth set makes it possible to "translate" a problem about compound statements into a problem about sets. It is also possible to go in the reverse direction. Given a problem about sets, think of the universal set as being a set of logical possibilities and think of a subset as being the truth set

of a statement. Hence we can "translate" a problem about sets into a problem about compound statements.

So far we have discussed only the truth sets assigned to compound statements involving $\vee$, $\wedge$, and $\sim$. All the other connectives can be defined in terms of these three basic ones, so that we can deduce what truth sets should be assigned to them. For example, we know that $p \rightarrow q$ is equivalent to $\sim p \vee q$ (see Figure 28 of Chapter I). Hence the truth set of $p \rightarrow q$ is the same as the truth set of $\sim p \vee q$, that is, it is $\tilde{P} \cup Q$. The Venn diagram for $p \rightarrow q$ is shown in Figure 6, where the shaded area is the truth set for the statement. Observe that the unshaded area in Figure 6 is the set $P - Q = P \cap \tilde{Q}$ which is the truth set of the statement $p \wedge \sim q$. Thus the shaded area is the set $\widetilde{(P - Q)} = \widetilde{P \cap \tilde{Q}}$ which is the truth set of the statement $\sim[p \wedge \sim q]$. We have thus discovered the fact $(p \rightarrow q)$, $(\sim p \vee q)$, and $\sim(p \wedge \sim q)$ are equivalent. It is always the case that two compound statements are equivalent if and only if they have the same truth sets. We also see that Venn diagrams can be used to discover relations between statements.

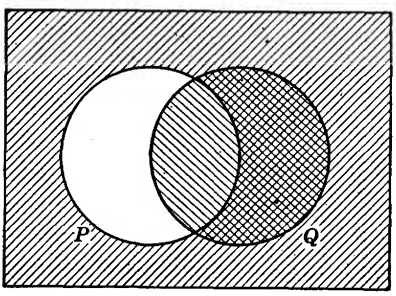


Figure 6

Suppose now that p is a statement that is logically true. What is its truth set? Now p is logically true if and only if it is true in every logically possible case, so that the truth set of p must be $\mathfrak{U}$. Similarly, if p is logically false, then it is false for every logically possible case, so that its truth set is the empty set $\mathcal{E}$.

Finally, let us consider the implication relation. Recall that p implies q if and only if the conditional $p \rightarrow q$ is logically true. But $p \rightarrow q$ is logically true if and only if its truth set is $\mathfrak{U}$, that is $\widetilde{(P - Q)} = \mathfrak{U}$, or $(P - Q) = \mathcal{E}$. From Figure 4 we see that if $P - Q$ is empty, then P is contained in Q. We shall symbolize the containing relation as follows: $P \subset Q$ means "P is a subset of Q." We conclude that $p \rightarrow q$ is logically true if and only if $P \subset Q$.

Let us briefly summarize the above discussion. To each statement

there corresponds a truth set. To each logical connective there corresponds a set operation. To each relation between statements there corresponds a relation between the truth sets. The truth sets of the statements $p \vee q$, $p \wedge q$, $\sim p$, and $p \rightarrow q$ are $P \cup Q$, $P \cap Q$, $\tilde{P}$, and $\widetilde{(P - Q)}$, respectively. Statement p is logically true if $P = \mathfrak{U}$ and logically false if $P = \mathcal{E}$. Statements p and q are equivalent if and only if $P = Q$, and p implies q if and only if $P \subset Q$.

Example 1. Prove by means of a Venn diagram that the statement $[p \vee (\sim p \vee q)]$ is logically true. The assigned set of this statement is $[P \cup (\tilde{P} \cup Q)]$, and its Venn diagram is shown in Figure 7. In that figure the set P is shaded vertically, and the set $\tilde{P} \cup Q$ is shaded horizontally. Their union is the entire shaded area which is $\mathfrak{U}$ so that the compound statement is logically true.

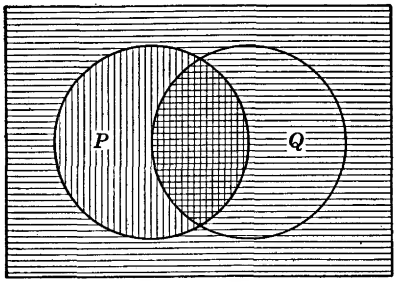


Figure 7

Example 2. Prove by means of Venn diagrams that $p \vee (q \wedge r)$ is equivalent to $(p \vee q) \wedge (p \vee r)$. The truth set of $p \vee (q \wedge r)$ is the entire shaded area of Figure 8, and the truth set of $(p \vee q) \wedge (p \vee r)$ is the doubly shaded area in Figure 9. Since these two sets are equal we see that the two statements are equivalent.

Example 3. Show by means of a Venn diagram that q implies $p \rightarrow q$. The truth set of $p \rightarrow q$ is the shaded area in Figure 6. Since this shaded area includes the set Q we see that q implies $p \rightarrow q$.

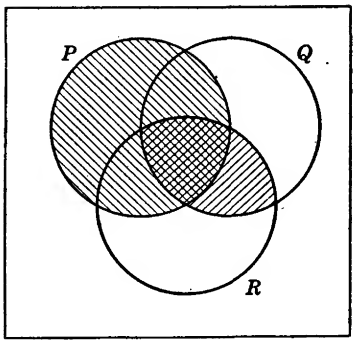


Figure 8

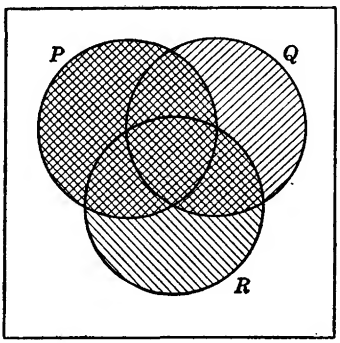


Figure 9

EXERCISES

Note: In Exercises 1, 2, and 3, find first the truth set of each statement.

1. Use Venn diagrams to test which of the following statements are logically true or logically false.

(a) $p \vee \sim p$.
(b) $p \wedge \sim p$.
(c) $p \vee (\sim p \wedge q)$.
(d) $p \rightarrow (q \rightarrow p)$.
(e) $p \wedge \sim(q \rightarrow p)$.

[*Ans.* (a), (d) logically true; (b), (e) logically false.]

2. Use Venn diagrams to test the following statements for equivalences.

(a) $p \vee \sim q$.
(b) $\sim(p \wedge q)$.
(c) $\sim(q \wedge \sim p)$.
(d) $p \rightarrow \sim q$.
(e) $\sim p \vee \sim q$.

[*Ans.* (a) and (c) equivalent; (b) and (d) and (e) equivalent.]

3. Use Venn diagrams for the following pairs of statements to test whether one implies the other.

(a) p; $p \wedge q$.
(b) $p \wedge \sim q$; $\sim p \rightarrow \sim q$.
(c) $p \rightarrow q$; $q \rightarrow p$.
(d) $p \wedge q$; $p \wedge \sim q$.

4. A pair of statements is said to be *inconsistent* if they cannot both be true. Devise a test for inconsistency.

5. Three or more statements are said to be inconsistent if they cannot all be true. What does this state about their truth sets?

6. In the following three compound statements (a) assign variables to the components, (b) bring the statements into symbolic form, (c) find the truth sets, and (d) test for consistency.

If this is a good course, then I will work hard in it.
If this is not a good course, then I shall get a bad grade in it.
I will not work hard, but I will get a good grade in this course.

[*Ans.* Inconsistent.]

Note: In Exercises 7-9 assign to each set a statement having it as a truth set.

7. Use truth tables to find which of the following sets are empty.
(a) $(P \cup Q) \cap (\tilde{P} \cup \tilde{Q})$.
(b) $(P \cap Q) \cap (\tilde{Q} \cap R)$.
(c) $(P \cap Q) - P$.
(d) $(P \cup R) \cap (\tilde{P} \cup \tilde{Q})$. [*Ans.* (b) and (c).]

8. Use truth tables to find out whether the following sets are all different.
(a) $P \cap (Q \cup R)$.
(b) $(R - Q) \cup (Q - R)$.
(c) $(R \cup Q) \cap (R \cap Q)$.
(d) $(P \cap Q) \cup (P \cap R)$.
(e) $(P \cap Q \cap \tilde{R}) \cup (P \cap \tilde{Q} \cap R) \cup (\tilde{P} \cap Q \cap \tilde{R}) \cup (\tilde{P} \cap \tilde{Q} \cap R)$.

9. Use truth tables for the following pairs of sets to test whether one is a subset of the other.
(a) P; $P \cap Q$.
(b) $P \cap \tilde{Q}$; $Q \cap \tilde{P}$.
(c) $P - Q$; $Q - P$.
(d) $\tilde{P} \cap \tilde{Q}$; $P \cup Q$.

10. Show, both by the use of truth tables and by the use of Venn diagrams, that $p \wedge (q \vee r)$ is equivalent to $(p \wedge q) \vee (p \wedge r)$.

*4. THE ABSTRACT LAWS OF SET OPERATIONS

The set operations which we have introduced obey some very simple abstract laws, which we shall list in this section. These laws can be proved by means of Venn diagrams or they can be translated into statements and checked by means of truth tables.

The abstract laws given below bear a close resemblance to the elementary algebraic laws with which the student is already familiar. The resemblance can be made even more striking by replacing $\cup$ by $+$ and $\cap$ by $\times$. For this reason, a set, its subsets and the laws of combination of subsets are considered an algebraic system, called a Boolean algebra—after the British mathematician George Boole who was the first person to study them from the algebraic point of view. Any other system obeying these laws, for example the system of compound statements studied in Chapter I, is also known as a Boolean algebra. We can study any of these systems from either the algebraic or the logical point of view.

Below are the basic laws of Boolean algebras. The proofs of these laws will be left as exercises.

The laws governing union and intersection:

A1. $A \cup A = A.$
A2. $A \cap A = A.$
A3. $A \cup B = B \cup A.$
A4. $A \cap B = B \cap A.$
A5. $A \cup (B \cup C) = (A \cup B) \cup C.$
A6. $A \cap (B \cap C) = (A \cap B) \cap C.$
A7. $A \cap (B \cup C) = (A \cap B) \cup (A \cap C).$
A8. $A \cup (B \cap C) = (A \cup B) \cap (A \cup C).$
A9. $A \cup \mathfrak{U} = \mathfrak{U}.$
A10. $A \cap \mathcal{E} = \mathcal{E}.$
A11. $A \cap \mathfrak{U} = A.$
A12. $A \cup \mathcal{E} = A.$

The laws governing complements:

B1. $\tilde{\tilde{A}} = A.$
B2. $A \cup \tilde{A} = \mathfrak{U}.$
B3. $A \cap \tilde{A} = \mathcal{E}.$
B4. $\widetilde{(A \cup B)} = \tilde{A} \cap \tilde{B}.$
B5. $\widetilde{(A \cap B)} = \tilde{A} \cup \tilde{B}.$
B6. $\tilde{\mathfrak{U}} = \mathcal{E}.$

The laws governing set-differences:

C1. $A - B = A \cap \tilde{B}.$

C2. $\mathfrak{U} - A = \tilde{A}$.
C3. $A - \mathfrak{U} = \varepsilon$.
C4. $A - \varepsilon = A$.
C5. $\varepsilon - A = \varepsilon$.
C6. $A - A = \varepsilon$.
C7. $(A - B) - C = A - (B \cup C)$.
C8. $A - (B - C) = (A - B) \cup (A \cap C)$.
C9. $A \cup (B - C) = (A \cup B) - (C - A)$.
C10. $A \cap (B - C) = (A \cap B) - (A \cap C)$.

EXERCISES

1. Test laws in the group A1-A12 by means of Venn diagrams.

2. "Translate" the A-laws into laws about compound statements. Test these by truth tables.

3. Test the laws in groups B and C by Venn diagrams.

4. "Translate" the B- and C-laws into laws about compound statements. Test these by means of truth tables.

5. Derive the following results from the 28 basic laws.
(a) $A = (A \cap B) \cup (A \cap \tilde{B})$.
(b) $A \cup B = (A \cap B) \cup (A \cap \tilde{B}) \cup (\tilde{A} \cap B)$.
(c) $A \cap (A \cup B) = A$.
(d) $A \cup (\tilde{A} \cap B) = A \cup B$.

6. From the A- and B-laws and from C1, derive C2-C6.

7. Use A1-A12 and C2-C10 to derive B1, B2, B3, and B6.

5. TWO-DIGIT NUMBER SYSTEMS

In the decimal number system one can write any number by using only the ten digits, 0, 1, 2, . . . , 9. Other number systems can be constructed which use either fewer or more digits. Probably the simplest number system is the *binary number system* which uses only the digits 0 and 1. We shall consider all the possible ways of forming number systems using only these two digits.

The two basic arithmetical operations are addition and multiplication. To understand any arithmetic system, it is necessary to know how to add or multiply any two digits together. Thus to understand the decimal system, we had to learn a multiplication table and an

addition table each of which had 100 entries. To understand the binary system we have to learn a multiplication and an addition table, each of which has only four entries. These are shown in Figure 10.

+	0	1
0	0	1
1	1	?

·	0	1
0	0	0
1	0	1

Figure 10

The multiplication table given there is completely determined by the two familiar rules that multiplying a number by zero gives zero, and multiplying a number by one leaves it unchanged. For addition we have only the rule that the addition of zero to a number does not change that number. The latter rule is sufficient to determine all but one of the entries in the addition table in Figure 10. We must still decide what shall be the sum $1 + 1$.

What are the possible ways in which we can complete the addition table? The only one digit numbers that we can use are 0 and 1, and these lead to interesting systems. Of the possible two-digit numbers we see that 00, 01 are the same as 0 and 1 and so do not give anything new. The number 11 or any greater number would introduce a "jump" in the table, hence the only other possibility is 10. The addition tables of these three different number systems are shown in Figure 11, and they all have the multiplication table shown in Figure 10. Each of these systems is interesting in itself as the interpretations below show.

Let us say that the parity of a positive integer is the fact of it being odd or even. Consider now the number system having the addition table in Figure 11(a) and let 0 represent "even" and 1 represent

+	0	1
0	0	1
1	1	0

(a)

+	0	1
0	0	1
1	1	1

(b)

+	0	1
0	0	1
1	1	10

(c)

Figure 11

"odd." The tables above now tell how the parity of a combination of two positive integers is related to the parity of each. Thus $0 \cdot 1 = 0$ tells us that the product of an even number and an odd number is

even, while $1 + 1 = 0$ tells us that the sum of two odd numbers is even, etc. Thus the first number system is that which we get from the arithmetic of the positive integers if we consider only the parity of numbers.

The second number system, which has the addition table in Figure 11(b), has an interpretation in terms of sets. Let 0 correspond to the empty set $\mathcal{E}$ and 1 correspond to the universal set $\mathcal{U}$. Let the addition of numbers correspond to the union of sets and let the multiplication of sets correspond to the intersection of sets. Then $0 \cdot 1 = 0$ tells us that $\mathcal{E} \cap \mathcal{U} = \mathcal{E}$ and $1 + 1 = 1$ tells us that $\mathcal{U} \cup \mathcal{U} = \mathcal{U}$. The student should give the interpretations for the other arithmetic computations possible for this number system.

Finally, the third number system, which has the addition table in Figure 11(c), is the so-called *binary number system*. Every ordinary integer can be written as a binary integer. Thus the binary 0 corresponds to the ordinary 0, and the binary unit 1 to the ordinary single unit. The binary number 10 means a "unit of higher order" and corresponds to the ordinary number two (not to ten). The binary number 100 then means two times two or four. In general, if $b_n b_{n-1} \ldots b_2 b_1 b_0$ is a binary number, where each digit is either 0 or 1, then the corresponding ordinary integer I is given by the formula

$$I = b_n \cdot 2^n + b_{n-1} \cdot 2^{n-1} + \ldots + b_2 \cdot 2^2 + b_1 \cdot 2 + b_0.$$

Thus the binary number 11001 corresponds to $2^4 + 2^3 + 1 = 16 + 8 + 1 = 25$. The table in Figure 12 shows some binary numbers and their integer equivalents.

Binary number	1	10	11	100	101	110	111	1000	10000	100000
Integer	1	2	3	4	5	6	7	8	16	32

Figure 12

Because electronic circuits are particularly well adapted to performing computations in the binary system, modern high-speed electronic computers are frequently constructed to work in the binary system.

Example. As an example of a computation, let us multiply 5 by 5 in the binary system. Since the binary equivalent of 5 is the number 101 the multiplication is done as follows.

$$\begin{array}{r} 101 \\ \underline{101} \\ 101 \\ 000 \\ \underline{101} \\ 11001 \end{array}$$

The answer is the binary number 11001, which we saw above was equivalent to the integer 25, the answer we expected to get.

EXERCISES

1. Complete the interpretations of the addition and multiplication tables, for the number systems representing (a) parity, (b) the sets $\mathfrak{U}$ and $\mathcal{E}$.

2. (a) What are the binary numbers corresponding to the integers 11, 52, 64, 98, 128, 144? [*Ans.* 1100010 corresponds to 98.]

(b) What integers correspond to the binary numbers 1111, 1010101, 1000000, 11011011? [*Ans.* 1010101 corresponds to 85.]

3. Carry out the following operations in the binary system. Check your answer.

(a) $29 + 20$.

(b) $9 \cdot 7$.

4. Of the laws listed below, which apply to each of the three systems?

(a) $x + y = y + x$.

(b) $x + x = x$.

(c) $x + x + x = x$.

5. Interpret $a + b$ to be the larger of the two numbers a and b, and $a \cdot b$ to be the smaller of the two. Write tables of "addition" and "multiplication" for the digits 0 and 1. Compare the result with the three systems given above. [*Ans.* Same as the $\mathfrak{U}$, $\mathcal{E}$ system.]

6. What do the laws A1-A10 of the last section tell us about the second number system established above?

7. The first number system above (about parity) can be interpreted to deal with the remainders of integers when divided by 2. An even number

leaves 0, an odd number leaves 1. Construct tables of addition and multiplication for remainders of integers when divided by 3. [*Hint:* These will be 3×3 tables.]

8. Given a set of four elements, suppose that we want to number its subsets. For a given subset write down a binary number as follows: The first digit is 1 if and only if the first element is in the subset, the second digit is 1 if and only if the second element is in the subset, etc. Prove that this assigns a unique number, from 0 to 15, to each subset.

9. In a multiple choice test the answers were numbered 1, 2, 4, and 8. The students were told that there might be no correct answer, or that one or more answers might be correct. They were told to *add* together the numbers of the correct answers (or to write 0 if no answer was correct).

(a) By using the result of Exercise 8, show that the resulting number gives the instructor all the information he wants.

(b) On a given question the correct sum was 7. Three students put down 4, 8, and 15, respectively. Which answer was most nearly correct? Which answer was worst? [*Ans.* 15 best, 8 worst.]

10. In the ternary number system numbers are expressed to the base 3, so that 201 in this system stands for $2 \cdot 3^2 + 0 \cdot 3 + 1 \cdot 1 = 19$.

(a) Write the numbers from 1 through 30 in this notation.

(b) Construct a table of addition and multiplication for the digits 0, 1, 2.

(c) Carry out the multiplication of $5 \cdot 5$ in this system. Check your answer.

11. Explain the meaning of the numeral '2907' in our ordinary (base 10) notation, in analogy to the formula I given for the binary system.

*6. VOTING COALITIONS

As an application of our set concepts we shall consider the significance of voting coalitions in voting bodies. Here the universal set is a set of human beings which form a decision-making body. For example, the universal set might be the members of a committee, or of a city council, or of a convention, or of the House of Representatives, etc. Each member can cast a certain number of votes. The decision as to whether or not a measure is passed can be decided by a simple majority rule, or $\frac{2}{3}$ majority, etc.

Suppose now that a subset of the members of the body forms a coalition in order to pass a measure. The question is whether or not they have enough votes to guarantee passage of the measure. If they

have enough votes to carry the measure, then we say they form a *winning coalition.* If the members *not* in the coalition can pass a measure of their own, then we say that the original coalition is a *losing coalition.* Finally, if the members of the coalition cannot carry their measure, and if the members not in the coalition cannot carry their measure, then the coalition is called a *blocking coalition.*

Let us restate these definitions in set-theoretic terms. A coalition C is winning if they have enough votes to carry an issue; coalition C is losing if the coalition $\tilde{C}$ is winning; and coalition C is blocking if neither C nor $\tilde{C}$ is a winning coalition.

The following facts are immediate consequences of these definitions. The complement of a winning coalition is a losing coalition. The complement of a losing coalition is a winning coalition. The complement of a blocking coalition is a blocking coalition.

Example 1. A committee consists of six men each having one vote. A simple majority vote will carry an issue. Then any coalition of four or more members is winning, any coalition with one or two members is losing, and any three-person coalition is blocking.

Example 2. Suppose in Example 1 one of the six members (say the chairman) is given the power to break ties. Then any three-person coalition of which he is a member is winning, while the other three-person coalitions are losing; hence there are no blocking coalitions. The other coalitions are as in Example 1.

Example 3. Let the universal set $\mathfrak{U}$ be the set $\{x, y, w, z\}$, where x and y each has one vote, w has two votes, and z has three votes. Suppose it takes five votes to carry a measure. Then the winning coalitions are: $\{z, w\}$, $\{z, x, y\}$, $\{z, w, x\}$, $\{z, w, y\}$, and $\mathfrak{U}$. The losing coalitions are the complements of these sets. Blocking coalitions are: $\{z\}$, $\{z, x\}$, $\{z, y\}$, $\{w, x\}$, $\{w, y\}$, and $\{w, x, y\}$.

The last example shows that it is not always necessary to list all members of a winning coalition. For example, if the coalition $\{z, w\}$ is winning, then it is obvious that the coalition $\{z, w, y\}$ is also winning. In general, if a coalition C is winning, then any other set that has C as a subset will also be winning. Thus we are led to the notion of a *minimal winning coalition.* A minimal winning coalition is a

winning coalition which contains no smaller winning coalition as a subset. Another way of stating this is that a minimal winning coalition is a winning coalition such that, if any member is lost from the coalition, then it ceases to be a winning coalition.

If we know the minimal winning coalitions, then we know everything that we need to know about the voting problem. The winning coalitions are all those sets that contain a minimal winning coalition, and the losing coalitions are the complements of the winning coalitions. All other sets are blocking coalitions.

In Example 1 the minimal winning coalitions are the sets containing four members. In Example 2 the minimal winning coalitions are the three-member coalitions that contain the tie-breaking member and the four-member coalitions that do not contain the tie-breaking member. The minimal winning coalitions in the third example are the sets $\{z, w\}$ and $\{z, x, y\}$.

Sometimes there are committee members who have special powers or lack of power. If a member can pass any measure he wishes without needing any one else to vote with him, then we call him a *dictator*. Thus member x is a dictator if and only if $\{x\}$ is a winning coalition. A somewhat weaker but still very powerful member is one who can by himself block any measure. If x is such a member, then we say that x has *veto power*. Thus x has veto power if and only if $\{x\}$ is a blocking coalition. Finally if x is not a member of any minimal winning coalition, we shall call him a *powerless* member. Thus x is powerless if and only if any winning coalition of which x is a member is a winning coalition without him.

Example 4. An interesting example of a decision-making body is the Security Council of the United Nations. The Security Council has eleven members consisting of the five permanent large-nation members called the Big Five, and six small nation members. In order that a measure be passed by the Council, seven members including all of the Big Five must vote for the measure. Thus the seven-member sets made up of the Big Five plus two small nations are the minimal winning coalitions. Then the losing coalitions are the sets that contain at most four small nations. The blocking coalitions are the sets that are neither winning nor losing. In particular, a unit set that contains one of the Big Five as a member is a blocking coalition.

This is the sense in which a Big Five member has a veto. [The possibility of "abstaining" is immaterial in the above discussion.]

EXERCISES

1. A committee has w, x, y, and z as members. Member w has two votes, the others have one vote each. List the winning, losing, and blocking coalitions.

2. A committee has n members, each with one vote. It takes a majority vote to carry an issue. What are the winning, losing, and blocking coalitions?

3. The Board of Estimate of New York City consists of eight members with voting strength as follows:

s.	Mayor	3 votes
t.	Controller	3
u.	Council President	3
v.	Brooklyn Borough President	2
w.	Manhattan Borough President	2
x.	Bronx Borough President	1
y.	Richmond Borough President	1
z.	Queens Borough President	1

A simple majority is needed to carry an issue. List the minimal winning coalitions. List the blocking coalitions. Do the same if we give the mayor the additional power to break ties.

4. A company has issued 100,000 shares of common stock and each share has one vote. How many shares must a stockholder have to be a dictator? How many to have a veto? [*Ans.* 50,001; 50,000.]

5. In Exercise 4, if the company requires a $\frac{2}{3}$ majority vote to carry an issue, how many shares must a stockholder have to be a dictator or to have a veto? [*Ans.* At least 66,667; at least 33,334.]

6. Prove that if a committee has a dictator as a member, then the remaining members are powerless.

7. We can define a maximal losing coalition in analogy to the minimal winning coalitions. What is the relation between the maximal losing and minimal winning coalitions? Do the maximal losing coalitions provide all relevant information?

8. Prove that any two minimal winning coalitions have at least one member in common.

9. Find all the blocking coalitions in the Security Council example.

10. Prove that if a man has veto power and if he together with any one other member can carry a measure, then the distribution of the remaining votes is irrelevant.

SUGGESTED READING

Birkhoff, G., and S. MacLane, *A Survey of Modern Algebra*, Macmillan, New York, 1953, Chapter XI.

Tarski, A., *Introduction to Logic*, Oxford, New York, 2d rev. ed., 1946, Chapter IV.

Mathematical Association of America, Committee on the Undergraduate Program, *Universal Mathematics*, Part II, Tulane Book Store, Tulane University, New Orleans, 1955, Chapter I.

Allendoerfer, C. B., and C. O. Oakley, *Principles of Mathematics*, McGraw-Hill, New York, 1955, Chapter V.

Johnstone, H. W., Jr., *Elementary Deductive Logic*, Crowell, New York, 1954, Part Three.

Chapter III

PARTITIONS AND COUNTING

1. PARTITIONS

The problem to be studied in this chapter can be most conveniently described in terms of partitions of a set. A *partition* of a set $\mathfrak{U}$ is a subdivision of the set into subsets that are disjoint and exhaustive, i.e., every element of $\mathfrak{U}$ must belong to one and only one of the subsets. The subsets A_i in the partition are called *cells*. Thus $[A_1, A_2, \ldots, A_r]$ is a partition of $\mathfrak{U}$ if two conditions are satisfied: (1) $A_i \cap A_j = \mathcal{E}$ if $i \neq j$ (the cells are disjoint) and (2) $A_1 \cup A_2 \cup \ldots \cup A_r = \mathfrak{U}$ (the cells are exhaustive).

Example 1. If $\mathfrak{U} = \{a, b, c, d, e\}$, then $[\{a, b\}, \{c, d, e\}]$ and $[\{b, c, e\}, \{a\}, \{d\}]$ and $[\{a\}, \{b\}, \{c\}, \{d\}, \{e\}]$ are three different partitions of $\mathfrak{U}$. The last is a partition into unit sets.

The process of going from a fine to a less fine analysis of a set of logical possibilities is actually carried out by means of a partition. For example, let us consider the logical possibilities for the first three games of the World Series if the Yankees play the Dodgers. We can list the possibilities in terms of the winner of each game as

$$\{\text{YYY, YYD, YDY, DYY, YDD, DDY, DYD, DDD}\}.$$

We form a partition by putting all the possibilities with the same number of wins for the Yankees in a single cell,

$$[\{\text{YYY}\}, \{\text{YYD, YDY, DYY}\}, \{\text{YDD, DDY, DYD}\}, \{\text{DDD}\}].$$

Thus, if we wish the possibilities to be Yankees win three games, win two, win one, win zero, then we are considering a less detailed analysis obtained from the former analysis by identifying the possibilities in each cell of the partition.

If $[A_1, A_2, \ldots, A_r]$ and $[B_1, B_2, B_3, \ldots, B_s]$ are two partitions of the same set $\mathfrak{U}$, we can obtain a new partition by considering the collection of all subsets of $\mathfrak{U}$ of the form $A_i \cap B_j$ (see Exercise 7). This new partition is called the *cross-partition* of the original two partitions.

Example 2. A common use of cross-partitions is in the problem of classification. For example, from the set $\mathfrak{U}$ of all life forms we can form the partition $[P, A]$ where P is the set of all plants and A is the set of all animals. We may also form the partition $[E, F]$ where E is the set of extinct life forms and F is the set of all existing life forms. The cross-partition

$$[P \cap E, \quad P \cap F, \quad A \cap E, \quad A \cap F]$$

gives a complete classification according to the two separate classifications.

Many of the examples with which we shall deal in the future will relate to processes which take place in stages. It will be convenient to use partitions and cross-partitions to represent the stages of the process. The graphical representation of such a process is, of course, a tree. For example, suppose that the process is such that we learn in succession the truth values of a series of statements relative to a given situation. If $\mathfrak{U}$ is the set of logical possibilities for the situation, and p is a statement relative to $\mathfrak{U}$, then the knowledge of the truth value of p amounts to knowing which cell of the partition $[P, \tilde{P}]$ contains the actual possibility. Recall that P is the truth set of p, and $\tilde{P}$ is the truth set of $\sim p$. Suppose now we discover the truth value of a second statement q. This information can again be described by a partition, namely, $[Q, \tilde{Q}]$. The two statements together give us information which can be represented by the cross-partition of these two partitions,

$$[P \cap Q, \quad P \cap \tilde{Q}, \quad \tilde{P} \cap Q, \quad \tilde{P} \cap \tilde{Q}].$$

That is, if we know the truth values of p and q, we also know which of the cells of this cross-partition contain the particular logical possibility describing the given situation. Conversely, if we knew which cell contained the possibility, we would know the truth values for the statements p and q.

The information obtained by the additional knowledge of the truth value of a third statement r, having a truth set R, can be represented by the cross-partition of the three partitions $[P, \tilde{P}]$, $[Q, \tilde{Q}]$, $[R, \tilde{R}]$. This cross-partition is

$$[P \cap Q \cap R,\ P \cap Q \cap \tilde{R},\ P \cap \tilde{Q} \cap R,\ \tilde{P} \cap Q \cap R,$$
$$P \cap \tilde{Q} \cap \tilde{R},\ \tilde{P} \cap Q \cap \tilde{R},\ \tilde{P} \cap \tilde{Q} \cap R,\ \tilde{P} \cap \tilde{Q} \cap \tilde{R}]$$

Notice that now we have the possibility narrowed down to being in one of $8 = 2^3$ possible cells. Similarly, if we knew the truth values of n statements, our partition would have 2^n cells.

If the set $\mathfrak{U}$ were to contain 2^{20} (approximately one million) logical possibilities, and if we were able to ask yes-no questions in such a way that the knowledge of the truth value of each question would cut the number of possibilities in half each time, then we could determine in 20 questions any given possibility in the set $\mathfrak{U}$. We could accomplish this kind of questioning, for example, if we had a list of all the possibilities and were allowed to ask "Is it in the first half?," and, if the answer is yes, then "Is it in the first one-fourth?," etc. In practice we ordinarily do not have such a list, and we can only approximate this procedure.

Example 3. In the familiar radio game of twenty questions it is not unusual for a contestant to try to carry out a partitioning of the above kind. For example, he may know that he is trying to guess a city. He might ask, "Is the city in North America?" and if the answer is yes, "Is it in the United States?" and if yes, "Is it west of the Mississippi?" and if no, "Is it in the New England states?" etc. Of course, the above procedure does not actually divide the possibilities exactly in half each time. The more nearly the answer to each question comes to dividing the possibilities in half, the more certain one can be of getting the answer in twenty questions, if there are at most a million possibilities.

EXERCISES

1. If $\mathfrak{U}$ is the set of integers from 1 to 6, find the cross-partitions of the following pairs of partitions.

(a) $[\{1, 2, 3\}, \{4, 5, 6\}]$ and $[\{1, 4\}, \{2, 3, 5, 6\}]$.

(b) $[\{1, 2, 3, 4, 5\}, \{6\}]$ and $[\{1, 3, 5\}, \{2, 6\}, \{4\}]$.

[*Ans.* (a) $\{1\}, \{2, 3\}, \{4\}, \{5, 6\}$.]

2. A coin is thrown three times. List the possibilities according to which side turns up each time. Give the partition formed by putting in the same cell all those possibilities for which the same number of heads occur.

3. Let p and q be two statements with truth set P and Q. What can be said about the cross-partition of $[P, \tilde{P}]$ and $[Q, \tilde{Q}]$ in the case that:

(a) p implies q. [*Ans.* $P \cap \tilde{Q} = \mathcal{E}$.]

(b) p is equivalent to q.

(c) p and q are inconsistent.

4. Consider the set of eight states consisting of Illinois, Colorado, Michigan, New York, Vermont, Texas, Alabama, and California.

(a) Show that in three "yes" or "no" questions one can identify any one of the eight states.

(b) Design a set of three "yes" or "no" questions which can be answered independently of each other and which will serve to identify any one of the states.

5. An unabridged dictionary contains about 600,000 words and 3000 pages. If a person chooses a word from such a dictionary, is it possible to identify this word by twenty "yes" or "no" questions? If so, describe the procedure that you would use and discuss the feasibility of the procedure.

[*Ans.* One solution is the following. Use 12 questions to locate the page, but then you may need 9 questions to locate the word.]

6. Mr. Jones has two parents, each of his parents had two parents, each of these had two parents, etc. Tracing a person's family tree back 40 generations (about 1000 years) gives Mr. Jones 2^{40} ancestors, which is more people than have been on the earth in the last 1000 years. What is wrong with this argument?

7. Let $[A_1, A_2, A_3]$ and $[B_1, B_2]$ be two partitions. Prove that the cross-partition of the two given partitions really is a partition, that is, it satisfies requirements (1) and (2) for partitions.

8. The cross-partition formed from the truth sets of n statements has 2^n cells. As seen in Chapter I, the truth table of a statement compounded from n statements has 2^n rows. What is the relationship between these two facts?

9. Let p and q be statements with truth sets P and Q. Form the partition $[P \cap Q, P \cap \tilde{Q}, \tilde{P} \cap Q, \tilde{P} \cap \tilde{Q}]$. State in each case below which of the cells must be empty in order to make the given statement a logically true statement.

(a) $p \rightarrow q$.
(b) $p \leftrightarrow q$.
(c) $p \vee \sim p$.
(d) p.

10. A partition $[A_1, A_2, \ldots, A_n]$ is said to be a *refinement* of the partition $[B_1, B_2, \ldots, B_m]$ if every A_i is a subset of some B_k. Show that a cross-partition of two partitions is a refinement of each of the partitions from which the cross-partition is formed.

11. Consider the partition of the people in the United States determined by classification according to states. The classification according to county determines a second partition. Show that this is a refinement of the first partition. Give a third partition which is different from each of these and is a refinement of both.

12. What can be said concerning the cross-partition of two partitions, one of which is a refinement of the other?

13. Given nine objects, of which it is known that eight have the same weight and one is heavier, show how, in two weighings with a pan balance, the heavy one can be identified.

14. Suppose that you are given thirteen objects, twelve of which are the same, but one is either heavier or lighter than the others. Show that, with three weighings using a pan balance, it is possible to identify the odd object. [A complete solution to this problem is given on page 42 of *Mathematical Snapshots*, second edition, by H. Steinhaus.]

15. A subject can be completely classified by introducing several simple subdivisions and taking their cross-partition. Thus, courses in college may be partitioned according to subject, level of advancement, number of students, hours per week, interests, etc. For each of the following subjects, introduce five or more partitions. How many cells are there in the complete classification (cross-partition) in each case?

(a) Detective stories. (b) Diseases.

*2. APPLICATIONS

Here we shall give three applications showing how partitions can be used to describe three different situations in mathematical terms. Examples like these will be more fully developed in later chapters.

Example 1. *A simple game.* Smith and Jones play the following game: Jones is to hold concealed in his hand either a \$1 or a \$2 bill. Smith is to guess which it is and gets the bill if he guesses correctly. We shall consider in a later chapter the amount Smith should pay to play the game in order to make it fair, but at the moment we are interested only in describing the possibilities for the play of the game. The game develops in two stages. First Jones chooses a \$1 or a \$2 bill, and secondly Smith guesses either 1 or 2. We can represent the ways that these stages can be carried out by a tree with four branches shown in Figure 1. The four ways that the game can be played are represented by the four paths of the tree which we denote by a_1, a_2, a_3, a_4.

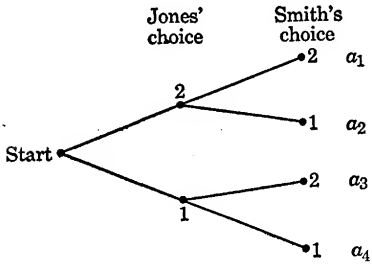


Figure 1

We can also represent the progress of the game by a sequence of three partitions,

$$\begin{array}{ll} \text{Start} & [\{a_1, a_2, a_3, a_4\}] \\ \text{Jones' Choice} & [\{a_1, a_2\}, \{a_3, a_4\}] \\ \text{Smith's Choice} & [\{a_1\}, \{a_2\}, \{a_3\}, \{a_4\}]. \end{array}$$

Notice that we have associated a partition with each level of the tree. A cell of the partition associated with a particular level contains all paths going through each branching point at this level.

We can also use partitions to indicate the amount of control which each player has on the outcome. The control of Jones can be indicated by the partition $[\{a_1, a_2\}, \{a_3, a_4\}]$. That is, Jones can determine which of the two cells of this partition will contain the play of the game. Similarly Smith can control the cell of the partition $[\{a_1, a_3\}, \{a_2, a_4\}]$ that will contain the play. The final partition is the cross-partition of these two partitions.

Example 2. *An example from psychology.* Suppose that a psychologist conducts the following experiment with a group of rats. Each rat is allowed in each trial to go through a T-maze of a type indicated

in Figure 2. If the rat chooses to go right, he will be fed, but if he goes left, he will not be fed. In other experiments different feeding schedules are used. For example, the food may always be put on the right, or it may be put on the right two-thirds (or any other fraction) of the time. The food might even be put each time on the side opposite the side to which the rat went the preceding trial. A psychologist is interested in predicting the behavior of the group of rats subjected to a sequence of such trials.

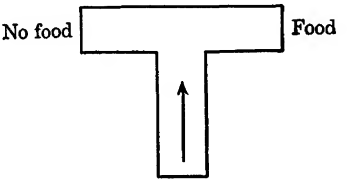


Figure 2

Consider an experiment of the above kind and denote by $\mathfrak{U}$ the set of all rats used in the experiment. After running each of the rats through the maze, we can form a partition by putting in one cell the rats which went right, and in the other those which went left. Thus each trial of the experiment determines a partition of $\mathfrak{U}$. The partition associated with the nth trial we denote by $[R_n, L_n]$, R_n being the set of rats which went right on the nth trial, and L_n those which went left. A psychologist would like to predict certain properties of the partitions after a large number of experiments. For example, questions he might ask are the following: if the food is always placed on the right, will R_n eventually become all, or almost all, of $\mathfrak{U}$? In other words, will the rats have "learned" to go right and be fed? What will happen if each rat is fed two-thirds of the time on the right and one-third on the left? What will happen if the experimenter does, on each trial, the opposite of what the rat did on the preceding trial?

Example 3. *Small group behavior.* Some sociologists study the behavior of a small group of people which has been given the job of jointly solving a problem. An example of this is a jury trying to decide the fate of a prisoner. Before a decision is reached, there is a good deal of discussion and argument among the members of the group, and experiments have been designed to study the role of each person in such a situation. For these experiments, observers record the name of the person making each remark together with the name of the person to whom the remark is directed. Sometimes the nature of the remark is recorded and also the time when it is made.

Consider an experiment of the above kind performed with four people, a, b, c, d. Let $\mathfrak{U}$ be the set of all remarks made. Form the partition $[S_a, S_b, S_c, S_d]$ of $\mathfrak{U}$ where S_a is the set of all remarks made by a, S_b the set of all remarks made by b, etc. Form also the partition $[T_a, T_b, T_c, T_d]$ of $\mathfrak{U}$ where T_a is the set of all remarks received by a, T_b is the set of all remarks received by b, etc. A sociologist is interested, for example, in the following question. Order the cells of the S partition according to the number of elements in each cell. If we do the same for the T partition, will the order be the same? That is, does the person who makes the most remarks also receive the most, and the one who makes the second most receive the second most, etc.?

A second problem is the following. Suppose that a partition $[U_1, U_2, U_3]$ of $\mathfrak{U}$ is made, where U_1 is the set of remarks made in a first interval of time, U_2 those made in a second interval of time, and U_3 those made in a third and final interval of time. Then if we form the cross-partition of this partition with each of the previous two partitions we will have a finer analysis which shows how the discussion changes in time. It might show, for example, that the discussion had changed from a three-way to a two-way discussion. It could also happen that eventually one person had made many remarks and received few. The nature of the partitions will of course depend upon the particular group of subjects and the particular experiment.

EXERCISES

1. Jones has two pennies and Smith has one. They agree to match pennies three times or until one of them has no pennies, whichever happens first. Draw a tree to represent the possible plays for the game. Show the progress of the game by a sequence of partitions. [*Ans.* There are four paths.]

2. In Example 2, what information is obtained from the cross-partition of the partitions $[R_1, L_1]$ and $[R_2, L_2]$?

3. Suppose that, in Example 2, the psychologist makes the following assumptions concerning the behavior of the rats subjected to a particular feeding schedule. For any particular trial, 80 per cent of the rats that went right on the previous experiment will go right on this trial, and 60 per cent of those that went left on the previous experiment will go right on this trial.

(a) If 50 per cent went right on the first trial, what per cent would the psychologist predict for the third trial? [*Ans.* 74 per cent.]

(b) If 75 per cent went right on the first trial, what per cent would he predict for the second? For the third? For the hundredth?

4. Construct a tree to represent the 16 possibilities for four runs of a rat through the T-maze. Check the possibilities in which the rat is successful at least three times if:

(a) The food is always on the right.

(b) The food is first on the right, then on the left, then on the right, and then on the left.

(c) The food is first on the right, then moved to the side opposite from the one that the rat went to the last time.

[*Ans.* There are five such paths.]

5. In Exercise 4, if the rat knows that one of the three possible methods of feeding is being used, how can he assure himself of three feedings in four tries? [*Ans.* Go right, left, right, right.]

6. In Example 3 on small group behavior, what information can be obtained from the cross-partition of $[S_a, S_b, S_c, S_d]$ and $[T_a, T_b, T_c, T_d]$?

7. Suppose that in Example 3, a partition $[V, \tilde{V}]$ of $\mathfrak{U}$ is made, where V is the set of all remarks that were made in the form of a question. What information can be obtained from the cross-partition of $[V, \tilde{V}]$ and $[S_a, S_b, S_c, S_d]$? From the cross-partition of $[V, \tilde{V}]$ and $[U_1, U_2, U_3]$?

8. Assume that every man is classified as a Republican or a Democrat. Let us start with a partition of the men of a given generation. Assume that we obtain a similar partition for their sons, and for their grandsons, etc., for several generations. What might be the questions a political scientist would wish to investigate, using these partitions?

9. Assume that in a given generation x men are Republicans and y are Democrats. Assume further that it is known that 20 per cent of the sons of Republicans are Democrats and 30 per cent of the sons of Democrats are Republicans in any generation. What conditions must x and y satisfy if there are to be the same number of Republicans in each generation? Assume that the total number of men remains at 50 million in each generation. Is there more than one choice for x and y? If not, what must x and y be?

[*Ans.* There are 30 million Republicans.]

10. Assume that there are 30 million Democrats and 20 million Republican men in the country. It is known that p per cent of the sons of Democrats are Republicans, and q per cent of the sons of Republicans are Democrats. If the total number of men remains 50 million, what condition must p and q satisfy so that the number in each party remains the same? Is there more than one choice of p and q?

3. THE NUMBER OF ELEMENTS IN A SET

The remainder of this chapter will be devoted to certain counting problems. For any set X we shall denote by $n(X)$ the number of elements in the set.

Suppose we know the number of elements in certain given sets and wish to know the number in other sets related to these by the operations of unions, intersections, and complementation. As an example, consider the following problem.

Suppose that we are told that 100 students take mathematics, and 150 students take economics. Can we then tell how many take either mathematics or economics? The answer is no, since clearly we would also need to know how many students take both courses. If we know that no student takes both courses, i.e., if we know that the two sets of students are disjoint, then the answer would be the sum of the two numbers or 250 students.

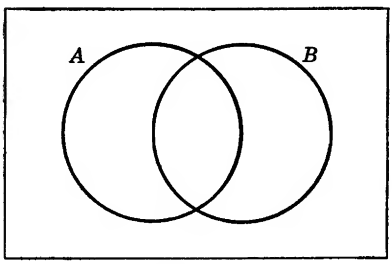


Figure 3

In general, if we are given disjoint sets A and B, then it is true that $n(A \cup B) = n(A) + n(B)$. Suppose now that A and B are not disjoint as shown in Figure 3. We can divide the set A into disjoint sets $A \cap \tilde{B}$ and $A \cap B$. Similarly we can divide B into the disjoint sets $\tilde{A} \cap B$ and $A \cap B$. Thus,

$$n(A) = n(A \cap \tilde{B}) + n(A \cap B)$$
$$n(B) = n(\tilde{A} \cap B) + n(A \cap B)$$

Adding these two equations, we obtain

$$n(A) + n(B) = n(A \cap \tilde{B}) + n(\tilde{A} \cap B) + 2n(A \cap B).$$

Since the sets $A \cap \tilde{B}$, $\tilde{A} \cap B$, and $A \cap B$ are disjoint sets whose union is $A \cup B$, we obtain the formula

$$n(A \cup B) = n(A) + n(B) - n(A \cap B)$$

which is valid for any two sets A and B.

Example 1. Let p and q be statements relative to a set $\mathfrak{U}$ of logical possibilities. Denote by P and Q the truth sets of these statements. The truth set of $p \vee q$ is $P \cup Q$ and the truth set of $p \wedge q$ is $P \cap Q$. Thus the above formula enables us to find the number of cases where $p \vee q$ is true if we know the number of cases for which p, q, and $p \wedge q$ are true.

Example 2. *More than two sets.* It is possible to derive formulas for the number of elements in a set which is the union of more than two sets (see Exercise 6), but usually it is easier to work with Venn diagrams. For example, suppose that the registrar of a school reports the following statistics about a group of 30 students:

19 take mathematics.
17 take music.
11 take history.
12 take mathematics and music.
7 take history and mathematics.
5 take music and history.
2 take mathematics, history, and music.

We draw the Venn diagram in Figure 4 and fill in the numbers for the number of elements in each subset working from the bottom of our list to the top. That is, since 2 students take all three courses, and 5 take music and history, then 3 take history and music but not mathematics, etc. Once the diagram is completed we can read off the number which take any combination of the courses. For example, the number which take history but not mathematics is $3 + 1 = 4$.

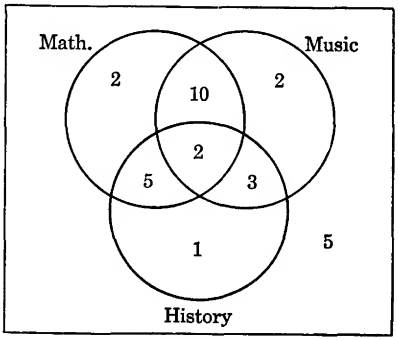


Figure 4

Example 3. *Cancer studies.* The following reasoning is often found in statistical studies on the effect of smoking on the incidence of lung cancer. Suppose a study has shown that the fraction of smokers

among those who have lung cancer is greater than the fraction of smokers among those who do not have lung cancer. It is then asserted that the fraction of smokers who have lung cancer is greater than the fraction of nonsmokers who have lung cancer. Let us examine this argument.

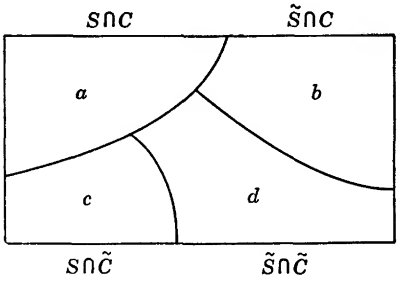


Figure 5

Let S be the set of all smokers in the population, and C be the set of all people with lung cancer. Let $a = n(S \cap C)$, $b = n(\tilde{S} \cap C)$, $c = n(S \cap \tilde{C})$, and $d = n(\tilde{S} \cap \tilde{C})$, as indicated in Figure 5. The fractions in which we are interested are

$$p_1 = \frac{a}{a+b}, \quad p_2 = \frac{c}{c+d}, \quad p_3 = \frac{a}{a+c}, \quad p_4 = \frac{b}{b+d},$$

where p_1 is the fraction of those with lung cancer that smoke, p_2 the fraction of those without lung cancer that smoke, p_3 the fraction of smokers who have lung cancer, and p_4 the fraction of nonsmokers who have cancer.

The argument above states that if $p_1 > p_2$, then $p_3 > p_4$. The hypothesis, $\frac{a}{a+b} > \frac{c}{c+d}$ is true if and only if $ac + ad > ac + bc$, that is, if and only if $ad > bc$. The conclusion $\frac{a}{a+c} > \frac{b}{b+d}$ is true if and only if $ab + ad > ab + bc$, that is, if and only if $ad > bc$. Thus the two statements $p_1 > p_2$ and $p_3 > p_4$ are in fact equivalent statements, so that the argument is valid.

EXERCISES

1. In Example 2 find:
   (a) The number of students that take mathematics but do not take history. [*Ans.* 12.]
   (b) The number that take exactly two of the three courses.
   (c) The number that take one or none of the courses.

2. In a chemistry class there are 20 students, and in a psychology class there are 30 students. Find the number in either the psychology class or the chemistry class if:

(a) The two classes meet at the same hour. [*Ans.* 50.]

(b) The two classes meet at different hours and 10 students are enrolled in both courses. [*Ans.* 40.]

3. If the truth set of a statement p has 10 elements, and the truth set of a statement q has 20 elements, find the number of elements in the truth set of $p \vee q$ if:

(a) p and q are inconsistent.

(b) p and q are consistent and there are 2 elements in the truth set of $p \wedge q$.

4. If p is a statement that is true in ten cases, and q is a statement that is true in five cases, find the number of cases that both p and q are true if $p \vee q$ is true in ten cases. What relation holds between p and q?

5. Assume that the incidence of lung cancer is 15 per 100,000, and that it is estimated that 75 per cent of those with lung cancer smoke and 60 per cent of those without lung cancer smoke. (These numbers are fictitious.) Estimate the fraction of smokers with lung cancer, and the fraction of non-smokers with lung cancer. [*Ans.* 18.75 and 9.375 per 100,000.]

6. Let A, B, and C be any three sets of a universal set $\mathfrak{U}$. Draw a Venn diagram and show that

$$n(A \cup B \cup C) = n(A) + n(B) + n(C) - n(A \cap B) - n(B \cap C) - n(A \cap C) + n(A \cap B \cap C).$$

7. Analyze the data given below and draw a Venn diagram like that in Figure 4. Assuming that every student in the school takes one of the courses, find the total number of students in the school.

(a)	(b)	
28	36	students take English.
23	23	students take French.
23	13	students take German.
12	6	students take English and French.
11	11	students take English and German.
8	4	students take French and German.
5	1	students take all three courses.

Comment on the result in (b).

8. Suppose that in a survey concerning the reading habits of students it is found that:

60 per cent read magazine A.
50 per cent read magazine B.
50 per cent read magazine C.
30 per cent read magazines A and B.
20 per cent read magazines B and C.
30 per cent read magazines A and C.
10 per cent read all three magazines.

(a) What per cent read exactly two magazines? [*Ans.* 50.]
(b) What per cent do not read any of the magazines? [*Ans.* 10.]

9. If p and q are equivalent statements and $n(P) = 10$, what is $n(P \cup Q)$?

10. If p implies q, prove that $n(P \cup \tilde{Q}) = n(P) + n(\tilde{Q})$.

11. On a transcontinental airliner, there are 9 boys, 5 American children, 9 men, 7 foreign boys, 14 Americans, 6 American males, and 7 foreign females. What is the number of people on the plane? [*Ans.* 33.]

4. PERMUTATIONS

We wish to consider here the number of ways in which a group of n different objects can be arranged. An arrangement of n different objects *in a given order* is called a *permutation* of the n objects. We consider first the case of three objects, a, b, and c. We can exhibit all possible permutations of these three objects as paths of a tree, as shown in Figure 6. Each path exhibits a possible permutation, and there are six such paths. We could also list these permutations as follows:

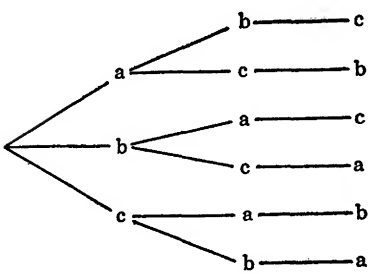


Figure 6

abc,	bca,
acb,	cab,
bac,	cba.

If we were to construct a similar tree for n objects, we would find that the number of paths could be found by multiplying together the numbers n, $n - 1$, $n - 2$, continuing down to the number 1. The number obtained in this way occurs so often that we give it

a symbol, namely $n!$, which is read "n factorial." Thus, for example, $3! = 3 \cdot 2 \cdot 1 = 6$, $4! = 4 \cdot 3 \cdot 2 \cdot 1 = 24$, etc. For reasons which will be clear later, we define $0! = 1$. Thus we can say *there are $n!$ different permutations of n distinct objects.*

Example 1. In the game of Scrabble, suppose there are seven lettered blocks from which we try to form a seven-letter word. If the seven letters are all different, we must consider $7! = 5040$ different orders.

Example 2. A quarterback has a sequence of ten plays. Suppose his coach instructs him to run through the ten-play sequence without repetition. How much freedom is left to the quarterback? He may choose any one of $10! = 3{,}628{,}800$ orders in which to call the plays.

Example 3. How many ways can n people be seated around a circular table? When this question is asked, it is usually understood that two arrangements are different only if at least one person has a different person next to him in the two arrangements. Consider then one person in a fixed position. There are $(n - 1)!$ ways in which the other people may be seated. We have now counted all the arrangements we wish to consider different. Why?

A general principle. There are many counting problems for which it is not possible to give a simple formula for the number of possible cases. In many of these the only way to find the number of cases is to draw a tree and count them (see Exercise 4). In some problems, the following general principle is useful.

If one thing can be done in exactly r different ways, for each of these a second thing can be done in exactly s different ways, for each of the first two, a third can be done in exactly t ways, etc., then the sequence of things can be done in the product of the numbers of ways in which the individual things can be done, i.e., $r \cdot s \cdot t \ldots$.

The validity of the above general principle can be established by thinking of a tree representing all the ways in which the sequence of things can be done. There would be r branches from the starting position. From the ends of each of these r branches there would be

s new branches, and from each of these t new branches, etc. The number of paths through the tree would be given by the product $r \cdot s \cdot t \ldots$.

Example 4. The number of permutations of n distinct objects is a special case of this principle. If we were to list all the possible permutations, there would be n possibilities for the first, for each of these $n - 1$ for the second, etc., until we came to the last object, and for which there is only one possibility. Thus there are $n(n-1)\ldots 1 = n!$ possibilities in all.

Example 5. If there are three roads from city x to city y and two roads from city y to city z, then there are $3 \cdot 2 = 6$ ways that a person can drive from city x to city z passing through city y.

Example 6. Suppose there are n applicants for a certain job. Three interviewers are asked independently to rank the applicants according to their suitability for the job. It is decided that an applicant will be hired if he is ranked first by at least two of the three interviewers. What fraction of the possible reports would lead to the acceptance of some candidate? We shall solve this problem by finding the fraction of the reports which do not lead to an acceptance and subtract this answer from 1. Frequently an indirect attack of this kind on a problem is easier than the direct approach. The total number of reports possible is $(n!)^3$ since each interviewer can rank the men in $n!$ different ways. If a particular report does not lead to the acceptance of a candidate, it must be true that each interviewer has put a different man in first place. This can be done in $n(n-1)(n-2)$ different ways by our general principle. For each possible first choices, there are $[(n-1)!]^3$ ways in which the remaining men can be ranked by the interviewers. Thus the number of reports which do not lead to acceptance is $n(n-1)(n-2)[(n-1)!]^3$. Dividing this number by $(n!)^3$ we obtain

$$\frac{(n-1)(n-2)}{n^2}$$

as the fraction of reports which fail to accept a candidate. The fraction which leads to acceptance is found by subtracting this fraction from 1 which gives

$$\frac{3n-2}{n^2}.$$

For the case of three applicants, we see that $\frac{7}{9}$ of the possibilities lead to acceptance. Here the procedure might be criticized on the grounds that even if the interviewers are completely ineffective and are essentially guessing there is a good chance that a candidate will be accepted on the basis of the reports. For n equal to ten, the fraction of acceptances is only .28, so that it is possible to attach more significance to the interviewers ratings, if they reach a decision.

EXERCISES

1. In how many ways can five people be lined up in a row for a group picture? In how many ways if it is desired to have three in the front row and two in the back row? [*Ans.* 120; 120.]

2. Assuming that a baseball team is determined by the players and the position each is playing, how many teams can be made from 13 players if:

(a) Each player can play any position?

(b) Two of the players can be used only as pitchers?

3. Grades of A, B, C, D, or E are assigned to a class of five students.

(a) How many ways may this be done, if no two students receive the same grade? [*Ans.* 120.]

(b) Two of the students are named Smith and Jones. How many ways can grades be assigned if no two students receive the same grade and Smith must receive a higher grade than Jones? [*Ans.* 60.]

(c) How many ways may grades be assigned if only grades of A and E are assigned? [*Ans.* 32.]

4. A certain club wishes to admit seven new members, four of whom are Republicans and three of whom are Democrats. Suppose the club wishes to admit them one at a time and in such a way that there are always more Republicans among the new members than there are Democrats. Draw a tree to represent all possible ways in which new members can be admitted, distinguishing members by their party only.

5. There are three different routes connecting city A to city B. How many ways can a round trip be made from A to B and back? How many ways if it is desired to take a different route on the way back? [*Ans.* 9; 6.]

6. How many different ways can a ten-question multiple-choice exam be answered if each question has three possibilities, a, b, and c? How many if no two consecutive answers are the same?

7. Modify Example 6 so that, to be accepted, an applicant must be first in two of the interviewers' rating and must be either first or second in the

third interviewers' rating. What fraction of the possible reports lead to acceptance in the case of three applicants? In the case of n?

[*Ans.* $\frac{4}{9}$; $4/n^2$.]

8. A town has 1240 registered Republicans. It is desired to contact each of these by phone to announce a meeting. A committee of r people devise a method of phoning s people each and asking each of these to call t new people. If the method is such that no person is called twice,

(a) How many people know about the meeting after the phoning?

(b) If the committee has 40 members and it is desired that all 1240 Republicans be informed of the meeting and that s and t should be the same, what should they be?

9. In the Scrabble example, suppose the letters are Q, Q, U, F, F, F, A. How many distinguishable arrangements are there for these seven letters? [*Ans.* 420.]

10. How many different necklaces can be made

(a) If seven different sized beads are available? [*Ans.* 360.]

(b) If six of the beads are the same size and one is larger? [*Ans.* 1.]

(c) If the beads are of two sizes, five of the smaller size and two of the larger size? [*Ans.* 3.]

11. Prove that two people in Columbus, Ohio, have the same initials.

12. Find the number of arrangements of the five symbols that can be distinguished. (The same letters with different subscripts indicate distinguishable objects.)

(a) A_1, A_2, B_1, B_2, B_3. [*Ans.* 120.]

(b) A, A, B_1, B_2, B_3. [*Ans.* 60.]

(c) A, A, B, B, B. [*Ans.* 10.]

13. Show that the number of distinguishable arrangements possible for n objects, n_1 of type 1, n_2 of type 2, etc., for r different types is

$$\frac{n!}{n_1!n_2!\ldots,n_r!}.$$

5. COUNTING PARTITIONS

Up to now we have not had occasion to consider the partitions $[(1, 2), (3, 4)]$ and $[(3, 4), (1, 2)]$ of the integers from 1 to 4 as being different partitions. Here it will be convenient to do so, and to indicate this distinction we shall use the term ordered partition. An *ordered partition with r cells* is a partition with r cells (some of which may be empty), with a particular order specified for the cells.

We are interested in counting the number of possible ordered partitions with r cells that can be formed from a set of n objects having a prescribed number of elements in each cell. We consider first a special case to illustrate the general procedure.

Suppose that we have eight students, A, B, C, D, E, F, G, and H, and we wish to assign these to three rooms, Room 1 which is a triple room, Room 2, a triple room, and Room 3, a double room. In how many different ways can the assignment be made? One way to assign the students is to put them in the rooms in the order in which they arrive, putting the first three in Room 1, the next three in Room 2, and the last two in Room 3. There are 8! ways in which the students can arrive, but not all of these lead to different assignments. We can represent the assignment corresponding to a particular order of arrival as follows,

$$|\text{BCA}|\text{DFE}|\text{HG}|.$$

In this case, B, C, and A are assigned to Room 1, D, F, and E to Room 2, and H and G to Room 3. Notice that orders of arrival which simply change the order within the rooms lead to the same assignment. The number of different orders of arrival which lead to the same assignment as the one above is the number of arrangements which differ from the given one only in that the arrangement within the rooms is different. There are $3!\cdot 3!\cdot 2!$ such orders of arrival, since we can arrange the three in Room 1 in 3! different ways, for each of these the ones in Room 2 in 3! different ways, and for each of these, the ones in Room 3 in 2! ways. Thus we can divide the 8! different orders of arrival into groups of $3!\cdot 3!\cdot 2!$ different orders such that all the orders of arrival in a single group lead to the same room assignment. Since there are $3!\cdot 3!\cdot 2!$ elements in each group and 8! elements altogether, there are $\dfrac{8!}{3!3!2!}$ groups, or this many different room assignments.

The same argument could be carried out for n elements and r rooms with n_1 in the first, n_2 in the second, etc. This would lead to the following result. Let $n_1, n_2, \ldots, n_r$ be nonnegative integers with $n_1 + n_2 + \ldots + n_r = n$. Then,

The number of ordered partitions with r cells $[A_1, A_2, A_3, \ldots, A_r]$ *of a set of n elements with* n_1 *in the first cell,* n_2 *in the second, etc. is*

$$\frac{n!}{n_1!n_2!\ldots n_r!}.$$

We shall denote this number by the symbol

$$\binom{n}{n_1,n_2,\ldots,n_r}.$$

The special case of two cells is particularly important. Here the problem can be stated equivalently as the problem of finding the number of subsets with r elements that can be chosen from a set of n elements. This is true because any choice defines a partition $[A, \tilde{A}]$, where A is the set of elements chosen and $\tilde{A}$ is the set of remaining elements. The number of such partitions is $\frac{n!}{r!(n-r)!}$ and hence this is also the number of subsets with r elements. Our notation $\binom{n}{r,n-r}$ for this case is shortened to $\binom{n}{r}$.

Notice that $\binom{n}{n-r}$ is the number of subsets with $n - r$ elements which can be chosen from n, which is the number of partitions of the form $[\tilde{A}, A]$ above. Clearly, this is the same as the number of $[A, \tilde{A}]$ partitions. Hence $\binom{n}{r} = \binom{n}{n-r}$.

Example 1. A college has scheduled six football games during a season. How many ways can the season end in two wins, three losses, and one tie? From each possible outcome of the season, we form a partition with three cells of the opposing teams. In the first cell we put the teams which our college defeats, in the second the teams to which our college loses, and in the third cell the teams which our college ties. There are $\binom{6}{2,3,1} = 60$ such partitions, and hence 60 ways in which the season can end with two wins, three losses, and one tie.

Example 2. In the game of bridge the hands N, E, S, and W determine a partition of the 52 cards having four cells each with thirteen

elements. Thus there are $\frac{52!}{13!13!13!13!}$ different bridge deals. This number is about $5.3645 \cdot 10^{28}$ or approximately 54 billion billion billion deals.

Example 3. The following example will be important in probability theory, which we take up in the next chapter. If a coin is thrown six times, there are 2^6 possibilities for the outcome of the six throws, since each throw can result in either a head or a tail. How many of these possibilities result in four heads and two tails? Each sequence of six heads and tails determines a two-cell partition of the numbers from one to six as follows: in the first cell put the numbers corresponding to throws which resulted in a head, and in the second put the numbers corresponding to throws which resulted in tails. We require that the first cell should contain four elements and the second two elements. Hence the number of the 2^6 possibilities which lead to four heads and two tails is the number of two-cell partitions of six elements which have four elements in the first cell and two in the second cell. The answer is $\binom{6}{4} = 15$. For n throws of a coin, a similar analysis shows that there are $\binom{n}{r}$ different sequences of H's and T's of length n which have exactly r heads and $n - r$ tails.

EXERCISES

1. Compute the following numbers.

(a) $\binom{7}{5}$ [*Ans.* 21.]

(b) $\binom{3}{2}$

(c) $\binom{7}{2}$

(d) $\binom{250}{249}$ [*Ans.* 250.]

(e) $\binom{5}{0}$

(f) $\binom{5}{1,2,2}$

(g) $\binom{4}{2,0,2}$ [*Ans.* 6.]

(h) $\binom{2}{1,1,1}$

2. Give an interpretation for $\binom{n}{0}$ and also for $\binom{n}{n}$. Can you now give a reason for making $0! = 1$?

3. How many ways can nine students be assigned to three triple rooms? How may ways if one particular pair of students refuse to room together? [*Ans.* 1680; 1260.]

4. A group of seven boys and ten girls attends a dance. If all the boys dance in a particular dance, how many possibilities are there for the girls who dance? For the girls who do not dance? How many possibilities are there for the girls who do not dance, if three of the girls are sure to be asked to dance?

5. Suppose that a course is given at three different hours. If fifteen students sign up for the course

(a) How many possibilities are there for the ways the students could distribute themselves in the classes? [*Ans.* 3^{15}.]

(b) How many of the ways would give the same number of students in each class? [*Ans.* 756,756.]

6. A college professor anticipates teaching the same course for the next 35 years. So not to become bored with his jokes, he decides to tell exactly three jokes every year and in no two years to tell exactly the same three jokes. What is the minimum number of jokes that will accomplish this? What is the minimum number if he determines never to tell the same joke twice?

7. How many ways can you answer a ten-question true-false exam, marking the same number of answers true as you do false? How many if it is desired to have no two consecutive answers the same?

8. From three Republicans and three Democrats, find the number of committees of three which can be formed,

(a) With no restrictions. [*Ans.* 20.]

(b) With three Republicans and no Democrats. [*Ans.* 1.]

(c) With two Republicans and one Democrat. [*Ans.* 9.]

(d) With one Republican and two Democrats. [*Ans.* 9.]

(e) With no Republicans and three Democrats. [*Ans.* 1.]

What is the relation between your answer in part (a) and the answers to the remaining four parts?

9. Problem 8 suggests that the following should be true.

$$\binom{2n}{n} = \binom{n}{0}\binom{n}{n} + \binom{n}{1}\binom{n}{n-1} + \binom{n}{2}\binom{n}{n-2} + \cdots + \binom{n}{n}\binom{n}{0}$$
$$= \binom{n}{0}^2 + \binom{n}{1}^2 + \cdots + \binom{n}{n}^2.$$

Show that it is true.

10. A student needs to choose two electives from six possible courses.
(a) How many ways can he make his choice? [*Ans.* 15.]
(b) How many ways can he choose if two of the courses meet at the same time? [*Ans.* 14.]
(c) How many ways can he choose if two of the courses meet at 10 o'clock, two at 11 o'clock, and there are no other conflicts among the courses? [*Ans.* 13.]

6. SOME PROPERTIES OF THE NUMBERS $\binom{n}{j}$

The numbers $\binom{n}{j}$ introduced in the previous section will play an important role in our future work. We give here some of the more important properties of these numbers.

A convenient way to obtain these numbers is given by the famous Pascal triangle, shown in Figure 7. To obtain the triangle we first

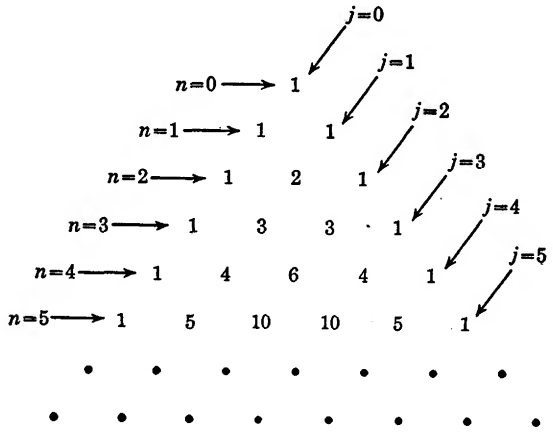


Figure 7

write the 1's down the sides. Any of the other numbers in the triangle has the property that it is the sum of the two adjacent numbers in the row just above. Thus the next row in the triangle is 1, 6, 15, 20, 15, 6, 1. To find the number $\binom{n}{j}$ we look in the row corresponding to the number n and see where the diagonal line corresponding to the

value of j intersects this row. For example, $\binom{4}{2} = 6$ is in the row marked $n = 4$ and on the diagonal marked $j = 2$.

The property of the numbers $\binom{n}{j}$ upon which the triangle is based is

$$\binom{n+1}{j} = \binom{n}{j-1} + \binom{n}{j}.$$

This fact can be verified directly (see Exercise 6), but the following argument is interesting in itself. The number $\binom{n+1}{j}$ is the number of subsets with j elements that can be formed from a set of $n+1$ elements. Select one of the $n+1$ elements, x. The $\binom{n+1}{j}$ subsets can be partitioned into those that contain x, and those that do not. The latter are subsets of j elements formed from n objects, and hence there are $\binom{n}{j}$ such subsets. The former are constructed by adding x to a subset of $j-1$ elements formed from n elements, and hence there are $\binom{n}{j-1}$ of them. Thus

$$\binom{n+1}{j} = \binom{n}{j-1} + \binom{n}{j}.$$

If we look again at the Pascal triangle, we observe that the numbers in a given row increase for a while, and then decrease. We can prove this fact in general by considering the ratio of two successive terms,

$$\frac{\binom{n}{j+1}}{\binom{n}{j}} = \frac{n!}{(j+1)!(n-j-1)!} \cdot \frac{j!(n-j)!}{n!} = \frac{n-j}{j+1}.$$

The numbers increase as long as the ratio is greater than 1, i.e., $n - j > j + 1$. This means that $j < \frac{1}{2}(n-1)$. We must distinguish the case of an even n from an odd n. For example, if $n = 10$, j must be less than $\frac{1}{2}(10-1) = 4.5$. Hence for j up to 4 the terms are increasing, from $j = 5$ on the terms decrease. For $n = 11$, j must be less than $\frac{1}{2}(11-1) = 5$. For $j = 5$, $(n-j)/(j+1) = 1$. Hence, up

to $j = 5$ the terms increase, then $\binom{11}{5} = \binom{11}{6}$, and then the terms decrease.

EXERCISES

1. Extend the Pascal triangle to $n = 16$. Save the result for later use.

2. Prove that

$$\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n} = 2^n,$$

using the fact that a set with n elements has 2^n subsets.

3. For a set of ten elements prove that there are more subsets with five elements than there are subsets with any other fixed number of elements.

4. Using the fact that $\binom{n}{r+1} = \frac{n-r}{r+1} \cdot \binom{n}{r}$, compute $\binom{30}{s}$ for $s =$ 1, 2, 3, 4 from the fact that $\binom{30}{0} = 1$. [*Ans.* 30; 435; 4060; 27,405.]

5. There are $\binom{52}{13}$ different possible bridge hands. Assume that a list is made showing all these hands, and that in this list the first card in every hand is crossed out. This leaves us with a list of twelve-card hands. Prove that at least two hands in the latter list contain exactly the same cards.

6. Prove that

$$\binom{n+1}{j} = \binom{n}{j-1} + \binom{n}{j},$$

using only the fact that

$$\binom{n}{j} = \frac{n!}{j!(n-j)!}.$$

7. Construct a triangle in the same way that the Pascal triangle was constructed, except that whenever you add two numbers, use the addition table in Chapter II, Figure 11(a). Construct the triangle for 16 rows. What does this triangle tell you about the numbers in the Pascal triangle? Use this result to check your triangle in Exercise 1.

8. In the triangle obtained in Exercise 7, what property do the rows 1, 2, 4, 8, and 16 have in common? What does this say about the numbers in the corresponding rows of the Pascal triangle? What would you predict for the terms in the 32nd row of the Pascal triangle?

9. For the following table state how one row is obtained from the preceding row and give the relation of this table to the Pascal triangle.

1	1	1	1	1	1	1
1	2	3	4	5	6	7
1	3	6	10	15	21	28
1	4	10	20	35	56	84
1	5	15	35	70	126	210
1	6	21	56	126	252	462
1	7	28	84	210	462	924

10. Referring to the table in Exercise 9, number the columns starting with 0, 1, 2, . . . and number the rows starting with 1, 2, 3, ... Let $f(n,r)$ be the element in the nth column and the rth row. The table was constructed by the rule

$$f(n,r) = f(n-1,r) + f(n,r-1)$$

for $n > 0$ and $r > 1$, and $f(n,1) = f(0,r) = 1$ for all n and r. Verify that

$$f(n,r) = \binom{n+r-1}{n}$$

satisfies these conditions and is in fact the only choice for $f(n,r)$ which will satisfy the conditions.

11. Consider a set $\{a, a, a\}$ of three objects which cannot be distinguished from one another. Then the ordered partitions with two cells which could be distinguished are

$$[\{a, a, a\}, \mathcal{E}]$$
$$[\{a, a\}, \{a\}]$$
$$[\{a\}, \{a, a\}]$$
$$[\mathcal{E}, \{a, a, a\}].$$

List all such ordered partitions with three cells. How many are there?

[*Ans.* 10.]

12. Let $f(n,r)$ be the number of distinguishable ordered partitions with r cells which can be formed from a set of n indistinguishable objects. Show that $f(n,r)$ satisfies the conditions

$$f(n,r) = f(n-1,r) + f(n,r-1)$$

for $n > 0$ and $r > 1$, and $f(n,1) = f(0,r) = 1$ for all n and r.

(*Hint:* Show that $f(n,r-1)$ is the number of partitions which have the last cell empty and $f(n-1,r)$ is the number which have at least one element in the last cell.)

13. Using the results of Exercises 10 and 12 show that the number of distinguishable ordered partitions with r cells which can be formed from a set of n indistinguishable objects is

$$\binom{n+r-1}{n}.$$

14. Assume that a mailman has seven letters to put in three mail boxes. How many ways can this be done if the letters are not distinguished?

[*Ans.* 36.]

15. By an ordered partition with r elements of n we mean a sequence of nonnegative integers, possibly some 0, written in a definite order, and having sum n. For example, $\{1, 0, 3\}$ and $\{3, 0, 1\}$ are two different ordered partitions with 3 elements of 4. Show that the number of ordered partitions with r elements of n is

$$\binom{n+r-1}{n}.$$

7. BINOMIAL AND MULTINOMIAL THEOREMS

It is sometimes necessary to expand products of the form $(x + y)^3$, $(x + 2y + 11z)^5$, etc. In this section we shall consider systematic ways of carrying out such expansions.

Consider first the special case $(x + y)^3$. We write this as

$$(x + y)^3 = (x + y)(x + y)(x + y).$$

To perform the multiplication, we choose either an x or a y from each of the three factors and multiply our choices together; we do this for all possible choices and add the results. We represent a particular set of choices by a two-cell partition of the numbers 1, 2, 3. In the first cell we put the numbers which correspond to factors from which we chose an x. In the second cell we put the numbers which correspond to factors from which we chose a y. For example, the partitions $[\{1, 3\}, \{2\}]$ correspond to a choice of x from the first and third factors and y from the second. The product so obtained is $xyx = x^2y$. The coefficient of x^2y in the expansion of $(x + y)^3$ will be the number of partitions which lead to a choice of two x's and one y. That is, the number of two-cell partitions of three elements with two elements in the first cell and one in the second, which is $\binom{3}{2} = 3$. More generally the coefficient of the term of the form x^iy^{3-i} will be $\binom{3}{j}$ for $j = 0, 1, 2, 3$. Thus we can write the desired expansion as

$$(x + y)^3 = \binom{3}{3}x^3 + \binom{3}{2}x^2y + \binom{3}{1}xy^2 + \binom{3}{0}y^3$$
$$= x^3 + 3x^2y + 3xy^2 + y^3.$$

The same argument carried out for the expansion $(x + y)^n$ leads to the binomial theorem of algebra.

Binomial theorem. The expansion of $(x + y)^n$ is given by

$$(x + y)^n = x^n + \binom{n}{n-1}x^{n-1}y + \binom{n}{n-2}x^{n-2}y^2 + \ldots + \binom{n}{1}xy^{n-1} + y^n.$$

Example 1. Let us find the expansion for $(a - 2b)^3$. To fit this into the binomial theorem, we think of x as being a and y as being $-2b$. Then we have

$$(a - 2b)^3 = a^3 + 3a^2(-2b) + 3a(-2b)^2 + (-2b)^3$$
$$= a^3 - 6a^2b + 12ab^2 - 8b^3.$$

We turn now to the problem of expanding the trinomial $(x + y + z)^3$. Again we write

$$(x + y + z)^3 = (x + y + z)(x + y + z)(x + y + z).$$

This time we choose either an x or y or z from each of the three factors. Our choice is now represented by a three-cell partition of the set of numbers $\{1, 2, 3\}$. The first cell has the numbers corresponding to factors from which we choose an x, the second cell the numbers corresponding to factors from which we choose a y, and the third those from which we choose a z. For example, the partition $[\{1, 3\}, \varepsilon, \{2\}]$ corresponds to a choice of x from the first and third factors, no y's, and a z from the second factor. The term obtained is $xzx = x^2z$. The coefficient of the term x^2z in the expansion is thus the number of three-cell partitions with two elements in the first cell, none in the second, and one in the third. There are $\binom{3}{2,0,1} = 3$ such partitions. In general the coefficient of the term of the form $x^ay^bz^c$ in the expan-

sion of $(x + y + z)^3$ will be $\binom{3}{a,b,c} = \frac{3!}{a!b!c!}$. Finding this way the coefficient for each possible a, b, and c we obtain

$$(x + y + z)^3 = x^3 + y^3 + z^3 + 3x^2y + 3xy^2 + 3yz^2 + 3y^2z + 3xz^2 + 3x^2z + 6xyz.$$

The same method can be carried out in general for finding the expansion of $(x_1 + x_2 + \ldots + x_r)^n$. From each factor we choose either an x_1, or x_2, or x_3, . . . , or x_r, form the product and add these products for all possible n choices. We will have r^n products, but many will be equal. A particular choice of one term from each factor determines an r-cell partition of the numbers from 1 to n. In the first cell we put the numbers of the factors from which we choose an x_1, in the second cell those from which we choose x_2, etc. A particular choice gives us a term of the form $x_1^{n_1} x_2^{n_2} \ldots x_r^{n_r}$ with $n_1 + n_2 + \ldots + n_r = n$. The corresponding partition has n_1 elements in the first cell, n_2 in the second, etc. For each such partition we obtain one such term. Hence the number of these terms which we obtain is the number of such partitions, which is

$$\binom{n}{n_1, n_2, \ldots, n_r} = \frac{n!}{n_1!n_2! \ldots n_r!}.$$

Thus we have the multinomial theorem.

Multinomial theorem. The expansion of $(x_1 + x_2 + \ldots + x_r)^n$ is found by adding all terms of the form

$$\binom{n}{n_1, n_2, \ldots, n_r} x_1^{n_1} x_2^{n_2} \ldots x_r^{n_r}$$

where $n_1 + n_2 + \ldots + n_r = n$.

EXERCISES

1. Expand by the binomial theorem
   (a) $(x + y)^4$.
   (b) $(1 + x)^5$.
   (c) $(x - y)^3$.
   (d) $(2x + a)^4$.
   (e) $(2x - 3y)^3$.
   (f) $(100 - 1)^5$.

2. Expand
   (a) $(x + y + z)^4$.
   (b) $(2x + y - z)^3$.
   (c) $(2 + 2 + 1)^3$. (Evaluate two ways.)

3. (a) Find the coefficient of the term $x^2y^3z^2$ in the expansion of $(x + y + z)^7$. [*Ans.* 210.]
   (b) Find the coefficient of the term $x^6y^3z^2$ in the expression $(x - 2y + 5z)^{11}$. [*Ans.* −924,000.]

4. Using the binomial theorem, prove that
   (a) $\binom{n}{0} + \binom{n}{1} + \binom{n}{2} + \ldots + \binom{n}{n} = 2^n$.
   (b) $\binom{n}{0} - \binom{n}{1} + \binom{n}{2} - \binom{n}{3} + \ldots \pm \binom{n}{n} = 0$.

5. Using an argument similar to the one in section 6, prove that
$$\binom{n+1}{i,j,k} = \binom{n}{i-1,j,k} + \binom{n}{i,j-1,k} + \binom{n}{i,j,k-1}.$$

6. Let $f(n,r)$ be the number of terms in the multinomial expansion of
$$(x_1 + x_2 + \ldots + x_r)^n;$$
and show that
$$f(n,r) = \binom{n+r-1}{n}.$$
(*Hint:* Show that the conditions of Section 6, Exercise 10 are satisfied by showing that $f(n,r-1)$ is the number of terms which do not have x_r and $f(n-1,r)$ is the number which do. Alternately, use Exercise 15 of Section 6 by showing that each term in the expansion determines an ordered sequence of r integers whose sum is n.)

7. How many terms are there in each of the expansions:
   (a) $(x + y + z)^6$? [*Ans.* 28.]
   (b) $(a + 2b + 5c + d)^4$? [*Ans.* 35.]
   (c) $(r + s + t + u + v)^6$? [*Ans.* 210.]

8. Prove that k^n is the sum of the numbers $\binom{n}{r_1,r_2,\ldots,r_k}$ for all choices of $r_1, r_2, \ldots, r_k$ such that $r_1 + r_2 + \ldots + r_k = n$.

*8. VOTING POWER

We return to the problem raised in Section 6 of Chapter II. Now we are interested not only in coalitions, but also in the power of indi-

vidual members. We will develop a numerical measure of voting power that was suggested by L. S. Shapley and M. Shubik. While the measure will be explained in detail below, for the reasons for choosing this particular measure the reader is referred to the original paper.

First of all we must realize that the number of votes a man controls is not in itself a good measure of his power. If x has three votes and y has one vote, it does not necessarily follow that x has three times the power that y has. Thus if the committee has just three members $\{x, y, z\}$ and z also has only one vote, then x is a dictator and y is powerless.

The basic idea of the power index is found in considering various alignments of the committee members on a number of issues. The n members are ordered $x_1, x_2, \ldots, x_n$ according to how likely they are to vote for the measure. If the measure is to carry, we must persuade x_1 and x_2 up to x_i to vote for it until we have a winning coalition. If $\{x_1, x_2, \ldots, x_i\}$ is a winning coalition but $\{x_1, x_2, \ldots, x_{i-1}\}$ is not winning, then x_i is the crucial member of the coalition. We must persuade him to vote for the measure, and he is the one hardest to persuade of the i necessary members. We call x_i the *pivot.*

For a purely mathematical measure of the power of a member we do not consider the views of the members. Rather we consider all possible ways that the members could be aligned on an issue, and see how often a given member would be the pivot. That means considering all permutations, and there will be $n!$ of them. In each permutation one member will be the pivot. The frequency with which a man is the pivot of an alignment is a good measure of his voting power.

DEFINITION. The voting power of a member of a committee is the number of alignments in which he is pivotal divided by the total number of alignments. (The total number of alignments, of course, is $n!$, for a committee of n members.)

Example 1. If all n members have one vote each, and it takes a majority vote to carry a measure, it is easy to see (by symmetry) that each member is pivot in $1/n$ of the alignments. Hence each member has power $= 1/n$. Let us illustrate this for $n = 3$. There are $3! = 6$ alignments. It takes two votes to carry a measure; hence the second member is always the pivot. The alignments are: 123, 132, 213, 231,

312, 321. The pivots are in **boldface.** Each member is pivot twice, hence has power $\frac{2}{6} = \frac{1}{3}$.

Example 2. Reconsider Chapter II, Section 6, Example 3 from this point of view. There are 24 permutations of the four members. We will list them, with the pivot in **boldface:**

w**x**yz	w**x**zy	w**y**xz	w**y**zx	w**z**xy	w**z**yx
x**w**yz	x**w**zy	x**y**wz	x**y**zw	x**z**wy	x**z**yw
y**x**wz	y**x**zw	y**w**xz	y**w**zx	y**z**xw	y**z**wx
z**x**yw	z**x**wy	z**y**xw	z**y**wx	z**w**xy	z**w**yx

We see that z has power of $\frac{14}{24}$, w has $\frac{6}{24}$, x and y have $\frac{2}{24}$ each. (Or, simplified, they have $\frac{7}{12}$, $\frac{3}{12}$, $\frac{1}{12}$, $\frac{1}{12}$ power, respectively.) We note that these ratios are much further apart than the ratio of votes which is 3:2:1:1. Here three votes are worth seven times as much as the single vote and more than twice as much as two votes.

Example 3. Reconsider Chapter II, Section 6, Example 4. By an analysis similar to the ones used so far (but too long to be included here) it can be shown that in the Security Council of the United Nations each of the Big Five has $\frac{76}{385}$ or approximately 0.197 power, while each of the small nations has approximately 0.002 power. This reproduces our intuitive feeling that, while the small nations in the Security Council are not powerless, nearly all the power is in the hands of the Big Five.

Example 4. In a committee of five each member has one vote, but the chairman has veto power. Hence the minimal winning coalitions are three-member coalitions including the chairman. There are $5! = 120$ permutations. The pivot cannot come before the chairman, since without the chairman we do not have a winning coalition. Hence, when the chairman is in place number 3, 4, or 5, he is the pivot. This happens in $\frac{3}{5}$ of the permutations. When he is in position 1 or 2, then the number 3 man is pivot. The number of permutations in which the chairman is in one of the first two positions and a given man is third is $2 \cdot 3! = 12$. Hence the chairman has power $\frac{3}{5}$, and each of the others has power $\frac{1}{10}$.

EXERCISES

1. A committee of three makes decisions by majority vote. Write out all permutations, and calculate the voting powers if the three members have:

(a) One vote each. [*Ans.* $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$.]

(b) One vote for two of them, two votes for the third. [*Ans.* $\frac{1}{6}, \frac{1}{6}, \frac{2}{3}$.]

(c) One vote for two of them, three votes for the third. [*Ans.* 0, 0, 1.]

(d) One, two, and three votes, respectively. [*Ans.* $\frac{1}{6}, \frac{1}{6}, \frac{2}{3}$.]

(e) Two votes each for two of them, and three votes for the third. [*Ans.* $\frac{1}{3}, \frac{1}{3}, \frac{1}{3}$.]

2. Prove that in any decision-making body the sum of the powers of the members is 1.

3. What is the power of a dictator? What is the power of a "powerless" member? Prove that your answers are correct.

4. A large company issued 100,000 shares. These are held by three stockholders, who have 50,000, 49,999, and 1 share, respectively. Calculate the powers of the three members. [*Ans.* $\frac{2}{3}, \frac{1}{6}, \frac{1}{6}$.]

5. A committee consists of 100 members having one vote each, plus a chairman who can break ties. Calculate the power distribution. (Do *not* try to write out all permutations!)

6. In Exercise 5, give the chairman a veto instead of the power to break ties. How does this change the power distribution?

[*Ans.* The chairman has power $\frac{50}{101}$.]

7. How are the powers in Exercise 1 changed if the committee requires a $\frac{3}{4}$ vote to carry a measure?

8. If in a committee of five, requiring majority decisions, each member has one vote, then each has power $\frac{1}{5}$. Now let us suppose that two members team up, and always vote the same way. Does this increase their power? (The best way to represent this situation is by allowing only those permutations in which these two members are next to each other.)

[*Ans.* Yes, the pair's power increases from .4 to .5.]

9. Given the votes that each member of a decision-making body controls, show that the minimal winning coalitions can be determined. If the minimal winning coalitions are known, show that the power of each member can be determined without knowing anything about the number of votes that each member controls.

10. Answer the following questions for a three-man committee:

(a) Find all possible sets of minimal winning coalitions.

(b) For each set of minimal winning coalitions find the distribution of voting power.
(c) Verify that the various distributions of power found in Exercises 1 and 7 are the only ones possible.

11. In Exercise 1, parts (a) and (e) have the same answer, and parts (b) and (d) and Exercise 4 also have the same answer. Use the results of Exercise 9 to find a reason for these coincidences.

12. Compute the voting power of one of the Big Five in the Security Council of the United Nations as follows:

(a) Show that for the nation to be pivotal it must be in the number 7 spot or later.
(b) Show that there are $\binom{6}{2}6!\,2!$ permutations in which the nation is in the number 7 spot.
(c) Find similar formulas for the number of permutations in which it is in the number 8, 9, 10, or 11 spot.
(d) Use this information to find the total number of permutations in which it is pivotal, and from this compute the power of the nation.

SUGGESTED READING

Mathematical Association of America, Committee on the Undergraduate Program, *Universal Mathematics*, Part II, Tulane Book Store, Tulane University, New Orleans, 1955, Chaps. III, IV.

Shapley, L. S., and M. Shubik, "A Method for Evaluating the Distribution of Power in a Committee System," *The American Political Science Review*, XLVIII (1954), pp. 787-792.

Whitworth, W. A., *Choice and Chance, with 1000 Exercises*, Stechert, New York, 1934.

Chapter IV

PROBABILITY THEORY

1. INTRODUCTION

We often hear statements of the following kind, "It is likely to rain today," "I have a fair chance of passing this course," "There is an even chance that a coin will come up heads," etc. In each case our statement refers to a situation in which we are not certain of the outcome, but we express some degree of confidence that our prediction will be verified. The theory of probability provides a mathematical framework for such assertions.

Consider an experiment whose outcome is not known. Suppose that someone makes an assertion p about the outcome of the experiment, and we want to assign a probability to p. When statement p is considered in isolation, we usually find no natural assignment of probabilities. Rather, we look for a method of assigning probabilities to all conceivable statements concerning the outcome of the experiment. At first this might seem to be a hopeless task, since there is no end to the statements we can make about the experiment. However we are aided by a basic principle:

Fundamental assumption. Any two equivalent statements will be assigned the same probability.

As long as there are a finite number of logical possibilities, there are only a finite number of truth sets, and hence the process of assigning probabilities is a finite one. We proceed in three steps: (1) we first determine $\mathfrak{U}$, the possibility set, that is, the set of all logical possibili-

ties, (2) to each subset X of $\mathfrak{U}$ we assign a number called the measure $m(X)$, (3) to each statement p we assign $m(P)$, the measure of its truth set, as a probability. The probability of statement p is denoted by $\Pr[p]$.

The first step, that of determining the set of logical possibilities, is one that we considered in the previous chapters. It is important to recall that there is no unique method for analyzing logical possibilities. In a given problem we may arrive at a very fine or a very rough analysis of possibilities, causing $\mathfrak{U}$ to have many or few elements.

Having chosen $\mathfrak{U}$, the next step is to assign a number to each subset X of $\mathfrak{U}$, which will in turn be taken to be the probability of any statement having truth set X. We do this in the following way.

Assignment of a measure. Assign a positive number (weight) to each element of $\mathfrak{U}$, so that the sum of the weights assigned is 1. Then the measure of a set is the sum of the weights of its elements. The measure of the set $\mathcal{E}$ is 0.

In applications of probability to scientific problems, the assignment of measures and the analysis of the logical possibilities may depend upon factual information and hence can best be done by the scientist making the application.

Once the weights are assigned, to find the probability of a particular statement we must find its truth set and find the sum of the weights assigned to elements of the truth set. This problem, which might seem easy, can often involve considerable mathematical difficulty. The development of techniques to solve this kind of problem is the main task of probability theory.

Example 1. An ordinary die is thrown. What is the probability that the number which turns up is less than 4? Here the possibility set is $\mathfrak{U} = \{1, 2, 3, 4, 5, 6\}$. The symmetry of the die suggests that each face should have the same probability of turning up. To make this so we assign weight $\frac{1}{6}$ to each of the outcomes. The truth set of the statement, "The number which turns up is less than 4," is $\{1, 2, 3\}$. Hence the probability of this statement is $\frac{3}{6} = \frac{1}{2}$, the sum of the weights of the elements in its truth set.

Example 2. A man attends a race involving three horses A, B, and C. He feels that A and B have the same chance of winning but

that A (and hence also B) is twice as likely to win as C is. What is the probability that A or C wins? We take as $\mathfrak{U}$ the set {A, B, C}. If we were to assign weight a to the outcome C, then we would assign weight $2a$ to each of the outcomes A and B. Since the sum of the weights must be 1, we have $2a + 2a + a = 1$, or $a = \frac{1}{5}$. Hence we assign weights $\frac{2}{5}$, $\frac{2}{5}$, $\frac{1}{5}$ to the outcomes A, B, and C, respectively. The truth set of the statement "Horse A or C wins" is {A, C}. The sum of the weights of the elements of this set is $\frac{2}{5} + \frac{1}{5} = \frac{3}{5}$. Hence the probability that A or C wins is $\frac{3}{5}$.

EXERCISES

1. Assume that there are n possibilities for the outcome of a given experiment. How should the weights be assigned if it is desired that all outcomes be assigned the same weight?

2. Let $\mathfrak{U} = \{a, b, c\}$. Assign weights to the three elements so that no two have the same weight, and find the measures of the eight subsets of $\mathfrak{U}$.

3. In an election Jones has probability $\frac{1}{2}$ of winning, Smith has probability $\frac{1}{3}$, and Black has probability $\frac{1}{6}$.

(a) Construct $\mathfrak{U}$.
(b) Assign weights.
(c) Find the measures of the eight subsets.
(d) Give a pair of nonequivalent predictions which have the same probability.

4. Give the possibility set $\mathfrak{U}$, for each of the following experiments.

(a) An election between candidates A and B is to take place.
(b) A number between 1 and 5 is chosen at random.
(c) A two-headed coin is thrown.
(d) A student is asked for the day of the year on which his birthday falls.

5. For which of the cases in Exercise 4 might it be appropriate to assign the same weight to each outcome?

6. Suppose that the following probabilities have been assigned to the possible results of putting a penny in a certain defective peanut-vending machine: The probability that nothing comes out is $\frac{1}{2}$. The probability that either you get your money back or you get peanuts (but not both) is $\frac{1}{3}$.

(a) What is the probability that you get your money back and also get peanuts? [*Ans.* $\frac{1}{6}$.]
(b) From the information given, is it possible to find the probability that you get peanuts? [*Ans.* No.]

7. A die is loaded in such a way that the probability of each face is proportional to the number of dots on that face. (For instance, a 6 is three times as probable as a 2.) What is the probability of getting an even number in one throw? [*Ans.* $\frac{4}{7}$.]

8. If a coin is thrown three times, list the eight possibilities for the outcomes of the three successive throws. A typical outcome can be written (HTH). Determine a probability measure by assigning an equal weight to each outcome. Find the probabilities of the following statements:

(r) The number of heads that occur is greater than the number of tails. [*Ans.* $\frac{1}{2}$.]

(s) Exactly two heads occur. [*Ans.* $\frac{3}{8}$.]

(t) The same side turns up on every throw. [*Ans.* $\frac{1}{4}$.]

9. For the statements given in Exercise 8, which of the following equalities are true?

(a) $\Pr[r \vee s] = \Pr[r] + \Pr[s]$

(b) $\Pr[s \vee t] = \Pr[s] + \Pr[t]$

(c) $\Pr[r \vee \sim r] = \Pr[r] + \Pr[\sim r]$

(d) $\Pr[r \vee t] = \Pr[r] + \Pr[t]$

10. Which of the following pairs of statements (see Exercise 8) are inconsistent? (Recall that two statements are inconsistent if their truth sets have no element in common.)

(a) r, s (c) r, $\sim r$

(b) s, t (d) r, t [*Ans.* (b) and (c).]

11. State a theorem suggested by Exercises 9 and 10.

2. PROPERTIES OF A PROBABILITY MEASURE

Before studying special probability measures, we shall consider some general properties of such measures which are useful in computations and in the general understanding of probability theory.

Three basic properties of a probability measure are:

(A) $m(X) = 0$ if and only if $X = \mathcal{E}$.

(B) $0 \leq m(X) \leq 1$ for any set X.

(C) For two sets X and Y,

$$m(X \cup Y) = m(X) + m(Y)$$

if and only if X and Y are disjoint, i.e., have no elements in common.

The proofs of properties (A) and (B) are left as an exercise (see Exercise 16). We shall prove (C).

We observe first that $m(X) + m(Y)$ is the sum of the weights of the elements of X added to the sum of the weights of Y. If X and Y are disjoint, then the weight of every element of $X \cup Y$ is added once and only once, and hence $m(X) + m(Y) = m(X \cup Y)$.

Assume now that X and Y are not disjoint. Here the weight of every element contained in both X and Y, i.e., in $X \cap Y$, is added twice in the sum $m(X) + m(Y)$. Thus this sum is greater than $m(X \cup Y)$ by an amount $m(X \cap Y)$. By (A) and (B), if $X \cap Y$ is not the empty set, then $m(X \cap Y) > 0$. Hence in this case we have $m(X) + m(Y) > m(X \cup Y)$. Thus if X and Y are not disjoint, the equality in (C) does not hold. Our proof shows that in general we have

(C′) For any two sets X and Y,

$$m(X \cup Y) = m(X) + m(Y) - m(X \cap Y)$$

Since the probabilities for statements are obtained directly from the probability measure $m(X)$, any property of $m(X)$ can be translated into a property about the probability of statements. For example, the above properties become, when expressed in terms of statements:

(a) $\Pr[p] = 0$ if and only if p is logically false.

(b) $0 \leq \Pr[p] \leq 1$ for any statement p.

(c) The equality

$$\Pr[p \vee q] = \Pr[p] + \Pr[q]$$

holds if and only if p and q are inconsistent.

(c′) For any two statements p and q,

$$\Pr[p \vee q] = \Pr[p] + \Pr[q] - \Pr[p \wedge q].$$

Another property of a probability measure which is often useful in computation is

(D) $m(\tilde{X}) = 1 - m(X)$,

or, in the language of statements,

(d) $\Pr[\sim p] = 1 - \Pr[p]$.

The proofs of (D) and (d) are left as an exercise (see Exercise 17).

It is important to observe that our probability measure assigns

probability 0 only to statements which are logically false, i.e., which are false for every logical possibility. Hence, a prediction that such a statement will be true is certain to be wrong. Similarly a statement is assigned probability 1 only if it is true in every case, i.e., logically true. Thus the prediction that a statement of this type will be true is certain to be correct. (While these properties of a probability measure seem quite natural, it is necessary, when dealing with infinite possibility sets, to weaken them slightly. We consider in this book only the finite possibility sets.)

We shall now discuss the interpretation of probabilities that are not 0 or 1. We shall give only some intuitive ideas that are commonly held concerning probabilities. While these ideas can be made mathematically more precise, we offer them here only as a guide to intuitive thinking.

Suppose that, relative to a given experiment, a statement has been assigned probability p. From this it is often inferred that if a sequence of such experiments is performed under identical conditions, the fraction of experiments which yield outcomes making the statement true would be approximately p. The mathematical version of this is the "law of large numbers" of probability theory (which will be treated in Section 10). In cases where there is no natural way to assign a probability measure, the probability of a statement is estimated experimentally. A sequence of experiments is performed and the fraction of the experiments which make the statement true is taken as the approximate probability for the statement.

A second and related interpretation of probabilities is concerned with betting. Suppose that a certain statement has been assigned probability p. We wish to offer a bet that the statement will in fact turn out to be true. We agree to give r dollars if the statement does not turn out to be true, provided that we receive s dollars if it does turn out to be true. What should r and s be to make the bet fair? If it were true that in a large number of such bets we would win s a fraction p of the times and lose r a fraction $1 - p$ of the time, then our average winning per bet would be $sp - r(1 - p)$. To make the bet fair we should make this average winning 0. This will be the case if $sp = r(1 - p)$ or if $r/s = p/(1 - p)$. Notice that this determines only the ratio of r and s. Such a ratio, written $r{:}s$, is said to give *odds* for the bet.

Example. Assume that a probability of $\frac{3}{4}$ has been assigned to a certain horse winning a race. Then the odds for a fair bet would be $\frac{3}{4}:\frac{1}{4}$. These odds could be equally well written as 3:1, 6:2 or 12:4, etc. A fair bet would be to agree to pay \$3 if the horse loses and receive \$1 if the horse wins. Another fair bet would be to pay \$6 if the horse loses and win \$2 if the horse wins.

EXERCISES

1. Let p and q be statements such that $\Pr[p \wedge q] = \frac{1}{4}$, $\Pr[\sim p] = \frac{1}{3}$, and $\Pr[q] = \frac{1}{2}$. What is $\Pr[p \vee q]$? [*Ans.* $\frac{11}{12}$.]

2. Using the result of Exercise 1, find $\Pr[\sim p \wedge \sim q]$.

3. Let p and q be statements such that $\Pr[p] = \frac{1}{2}$ and $\Pr[q] = \frac{2}{3}$. Are p and q consistent? [*Ans.* Yes.]

4. Show that, if $\Pr[p] + \Pr[q] > 1$, then p and q are consistent.

5. A student is worried about his grades in English and Art. He estimates that the probability of passing English is .4, that he will pass at least one course with probability .6, but that he has only probability .1 of passing both courses. What is the probability that he will pass Art? [*Ans.* .3.]

6. Given that a school has grades A, B, C, D, and F, and that a student has probability .9 of passing a course, and .6 of getting a grade lower than B, what is the probability that he will get a C or D? [*Ans.* $\frac{1}{2}$.]

7. What odds should a person give on a bet that a six will turn up when a die is thrown?

8. Referring to Example 2 of Section 1, what odds should the man be willing to give for a bet that either A or B will come in first?

9. Prove that if the odds relative to a given statement are $r:s$, then the probability that the statement will be true is $r/(r + s)$.

10. Using the result of Exercise 9 and the definition of "odds," show that if the odds are $r:s$ that a statement is true, then the odds are $s:r$ that it is false.

11. A man is willing to give 5:4 odds that the Dodgers will win the World Series. What must the probability of a Dodger victory be for this to be a fair bet? [*Ans.* $\frac{5}{9}$.]

12. A man has found through long experience that if he washes his car it rains the next day 85 per cent of the time. What odds should he give that this will occur next time?

13. A man offers 1:3 odds that A will occur, 1:2 odds that B will occur. He knows that A and B cannot both occur. What odds should he give that A or B will occur? [*Ans.* 7:5.]

14. A man offers 3:1 odds that A will occur, 2:1 odds that B will occur. He knows that A and B cannot both occur. What odds should he give that A or B will occur?

15. Show from the definition of a probability measure that $m(X) = 1$ if and only if $X = \mathfrak{U}$.

16. Show from the definition of a probability measure that properties (A), (B) of the text are true.

17. Prove property (D) of the text. Why does property (d) follow from this property?

18. Prove that if R, S, and T are three sets that have no element in common,

$$m(R \cup S \cup T) = m(R) + m(S) + m(T).$$

19. If X and Y are two sets such that X is a subset of Y, prove that $m(X) \leq m(Y)$.

20. If p and q are two statements such that p implies q, prove that $\Pr[p] \leq \Pr[q]$.

21. Suppose that you are given n statements and each has been assigned a probability less than or equal to r. Prove that the probability of the disjunction of these statements is less than or equal to nr.

22. The following is an alternative proof of property (C′) of the text. Give a reason for each step.

(a) $X \cup Y = (X \cap \tilde{Y}) \cup (X \cap Y) \cup (Y \cap \tilde{X})$.
(b) $m(X \cup Y) = m(X \cap \tilde{Y}) + m(X \cap Y) + m(\tilde{X} \cap Y)$.
(c) $m(X \cup Y) = m(X) + m(Y) - m(X \cap Y)$.

23. If X, Y, and Z are any three sets, prove that, for any probability measure,

$$m(X \cup Y \cup Z) = m(X) + m(Y) + m(Z) - m(X \cap Y) - m(Y \cap Z) - m(X \cap Z) + m(X \cap Y \cap Z).$$

24. Translate the result of Exercise 23 into a result concerning three statements p, q, and r.

25. A man offers to bet "dollars to doughnuts" that a certain event will take place. Assuming that a doughnut costs a nickel, what must the probability of the event be for this to be a fair bet? [*Ans.* $\frac{20}{21}$.]

3. THE EQUIPROBABLE MEASURE

We have already seen several examples where it was natural to assign the same weight to all possibilities in determining the appropriate probability measure. The probability measure determined in this manner is called the *equiprobable measure.* The measure of sets in the case of the equiprobable measure has a very simple form. In fact, if $\mathfrak{U}$ has n elements and if the equiprobable measure has been assigned, then for any set X, $m(X)$ is r/n, where r is the number of elements in the set X. This is true since the weight of each element in X is $1/n$, and hence the sum of the weights of elements of X is r/n.

The particularly simple form of the equiprobable measure makes it easy to work with. In view of this it is important to observe that a particular choice for the set of possibilities in a given situation may lead to the equiprobable measure, while some other choice will not. For example, consider the case of two throws of an ordinary coin. Suppose that we are interested in statements about the number of heads which occur. If we take for the possibility set the set $\mathfrak{U} = \{\text{HH, HT, TH, TT}\}$ then it is reasonable to assign the same weight to each outcome, and we are led to the equiprobable measure. If, on the other hand, we were to take as possible outcomes the set $\mathfrak{U} = \{\text{no H, one H, two H}\}$, it would not be natural to assign the same weight to each outcome, since one head can occur in two different ways, while each of the other possibilities can occur in only one way.

Example 1. Suppose that we throw two ordinary dice. Each die can turn up a number from 1 to 6; hence there are $6 \cdot 6$ possibilities. We assign weight $\frac{1}{36}$ to each possibility. A prediction that is true in j cases will then have probability $j/36$. For example, "The sum of the dice is 5," will be true if we get $1 + 4$, $2 + 3$, $3 + 2$, or $4 + 1$. Hence the probability that the sum of the dice is 5 is $\frac{4}{36} = \frac{1}{9}$. The sum can be 12 in only one way, $6 + 6$. Hence the probability that the sum is 12 is $\frac{1}{36}$.

Example 2. Suppose that two cards are drawn successively from a deck of cards. What is the probability that both are hearts? There are 52 possibilities for the first card, and for each of these there are 51 possibilities for the second. Hence there are $52 \cdot 51$ possibilities for the result of the two draws. We assign the equiprobable measure.

The statement "The two cards are hearts" is true in $13 \cdot 12$ of the $52 \cdot 51$ possibilities. Hence the probability of this statement is $13 \cdot 12/52 \cdot 51 = \frac{1}{17}$.

Example 3. Assume that, on the basis of a predictive index applied to students A, B, and C when entering college, it is predicted that after four years of college the scholastic record of A will be the highest, C the second highest, and B the lowest of the three. Suppose, in fact, that these predictions turn out to be exactly correct. If the predictive index has no merit at all and hence the predictions amount simply to guessing, what is the probability that such a prediction will be correct? There are $3! = 6$ orders in which the men might finish. If the predictions were really just guessing, then we would assign an equal weight to each of the six outcomes. In this case the probability that a particular prediction is true is $\frac{1}{6}$. Since this probability is reasonably large, we would hesitate to conclude that the predictive index is in fact useful, on the basis of this one experiment. Suppose, on the other hand, it predicted the order of six men correctly. Then a similar analysis would show that, by guessing, the probability is $\frac{1}{6!} = \frac{1}{720}$ that such a prediction would be correct. Hence, we might conclude here that there is strong evidence that the index has some merit.

EXERCISES

1. A letter is chosen at random from the word "random." What is the probability that it is an n? That it is a vowel? [*Ans.* $\frac{1}{6}$; $\frac{1}{3}$.]

2. An integer between 3 and 12 inclusive is chosen at random. What is the probability that it is an even number? That it is even and divisible by three?

3. A card is drawn at random from a pack of playing cards.

(a) What is the probability that it is either a heart or the king of clubs? [*Ans.* $\frac{7}{26}$.]

(b) What is the probability that it is either the queen of hearts or an honor card (i.e., ten, jack, queen, king, or ace)? [*Ans.* $\frac{5}{13}$.]

4. A word is chosen at random from the set of words $\mathfrak{U} =$ {men, bird, ball, field, book}. Let p, q, and r be the statements:

p: The word has two vowels.

q: The first letter of the word is b.

r: The word rhymes with cook.

Find the probability of the following statements:

(a) p.
(b) q.
(c) r.
(d) $p \wedge q$.
(e) $(p \vee q) \wedge \sim r$.
(f) $p \rightarrow q$. [*Ans.* $\frac{4}{5}$.]

5. A single die is thrown. Find the probability that

(a) An odd number turns up.
(b) The number which turns up is greater than two.
(c) A seven turns up.

6. In the primary voting example of Chapter II, Section 1, assume that all 36 possibilities in the elections are equally likely. Find

(a) The probability that candidate A wins more states than either of his rivals. [*Ans.* $\frac{7}{18}$.]
(b) That all the states are won by the same candidate. [*Ans.* $\frac{1}{36}$.]
(c) That every state is won by a different candidate. [*Ans.* 0.]

7. A single die is thrown twice. What value for the sum of the two outcomes has the highest probability? What value or values of the sum has the lowest probability of occurring?

8. Two boys and two girls are placed at random in a row for a picture. What is the probability that the boys and girls alternate in the picture? [*Ans.* $\frac{1}{3}$.]

9. A certain college has 500 students and it is known that:

300 read French.
200 read German.
50 read Russian.
20 read French and Russian.
30 read German and Russian.
20 read German and French.
10 read all three languages.

If a student is chosen at random from the school, what is the probability that the student:

(a) Reads two and only two languages?
(b) Reads at least one language?

10. Suppose that three people enter a restaurant which has a row of six seats. If they choose their seats at random, what is the probability that they sit with no seats between them? What is the probability that there is at least one empty seat between any two of them?

11. Find the probability of obtaining each of the following poker hands. (A poker hand is a set of five cards chosen at random from a deck of 52 cards.)

(a) Royal flush (ten, jack, queen, king, ace in a single suit.)
[*Ans.* $4/\binom{52}{5} = .0000015$.]

(b) Straight flush (five in a sequence in a single suit, but not a royal flush). [*Ans.* $(40 - 4)/\binom{52}{5} = .000014$.]

(c) Four of a kind (four cards of the same face value).
[*Ans.* $624/\binom{52}{5} = .00024$.]

(d) Full house (one pair and one triple of the same face value).
[*Ans.* $3744/\binom{52}{5} = .0014$.]

(e) Flush (five cards in a single suit but not a straight or royal flush).
[*Ans.* $(5148 - 40)/\binom{52}{5} = .0020$.]

(f) Straight (five cards in a row, not all of the same suit).
[*Ans.* $(10{,}240 - 40)/\binom{52}{5} = .0039$.]

(g) Straight or better. [*Ans.* .0076.]

12. If ten people are seated at a circular table at random, what is the probability that a particular pair of people are seated next to each other? [*Ans.* $\frac{2}{9}$.]

13. A room contains a group of n people who are wearing badges numbered from 1 to n. If two people are selected at random, what is the probability that the larger badge number is a 3? Answer this problem assuming that $n = 5, 4, 3, 2$. [*Ans.* $\frac{1}{5}; \frac{1}{3}; \frac{2}{3}; 0$.]

14. In Exercise 13, suppose that we observe two men leaving the room and that the larger of their badge numbers is 3. What might we guess as to the number of people in the room?

15. Find the probability that a bridge hand will have suits of:

(a) 5, 4, 3, and 1 cards. [*Ans.* $\dfrac{4!\binom{13}{5}\binom{13}{4}\binom{13}{3}\binom{13}{1}}{\binom{52}{13}} \cong .129$.]

(b) 6, 4, 2, and 1 cards. [*Ans.* .047.]

(c) 4, 4, 3, and 2 cards. [*Ans.* .216.]

(d) 4, 3, 3, and 3 cards. [*Ans.* .105.]

16. There are $\binom{52}{13} = 6.35 \times 10^{11}$ possible bridge hands. Find the probability that a bridge hand dealt at random will be all of one suit. Estimate *roughly* the number of bridge hands dealt in the entire country in a year. Is it likely that a hand of all one suit will occur sometime during the year in the United States?

4. TWO NONINTUITIVE EXAMPLES

There are occasions in probability theory when one finds a problem for which the answer, based on probability theory, is not at all in

agreement with one's intuition. It is usually possible to arrange a few wagers that will bring one's intuition into line with the mathematical theory. A particularly good example of this is provided by the matching birthdays problem.

Assume that we have a room with r people in it and we propose the bet that there are at least two people in the room having the same birthday, i.e., the same month and day of the year. We ask for the value of r which will make this a fair bet. Few people would be willing to bet even money on this wager unless there were at least 100 people in the room. Most people would suggest 150 as a reasonable number. However, we shall see that with 150 people the odds are approximately 4,500,000,000,000,000 to 1 in favor of two people having the same birthday, and that one should be willing to bet even money with as few as 23 people in the room.

Let us first find the probability that in a room with r people, no two have the same birthday. There are 365 possibilities for each person's birthday (neglecting February 29). There are then 365^r possibilities for the birthdays of r people. We assume that all these

Number of people in the room	Probability of at least two with same birthday	Approximate odds for a fair bet
5	.027	
10	.117	
15	.253	
20	.411	70:100
21	.444	80:100
22	.476	91:100
23	.507	103:100
24	.538	117:100
25	.569	132:100
30	.706	242:100
40	.891	819:100
50	.970	33:1
60	.994	169:1
70		1,200:1
80		12,000:1
90		160,000:1
100		3,300,000:1
125		31,000,000,000:1
150		4,500,000,000,000,000:1

Figure 1

possibilities are equally likely. To find the probability that no two have the same birthday we must find the number of possibilities for the birthdays which have *no* day represented twice. The first person can have any of 365 days for his birthday. For each of these, if the second person is to have a different birthday, there are only 364 possibilities for his birthday. For the third man, there are 363 possibilities if he is to have a different birthday than the first two, etc. Thus the probability that no two people have the same birthday in a group of r people is

$$q_r = \frac{365 \cdot 364 \cdot \ldots \cdot (365 - r + 1)}{365^r}.$$

The probability that at least two people have the same birthday is then $p_r = 1 - q_r$. In Figure 1 the values of p_r and the odds for a fair bet, $p_r{:}(1 - p_r)$, are given for several values of r.

We consider now a second problem in which intuition does not lead to the correct answer. We have seen that there are $n!$ permutations of the numbers from 1 to n. Let us consider a rearrangement of these numbers as the operation of placing each of the numbers in one of n boxes or *positions* (one number to a position). The positions are assumed to be numbered in serial order. We shall say that the ith number is unchanged by the permutation if, after the rearrangement, number i is still in the ith position. For example, if we consider the permutations of the numbers 1, 2, and 3, then the permutation 123 leaves *all* numbers fixed, the permutation 213 leaves *one* number fixed, and the permutations 312 and 231 leave *no* numbers fixed. It is obviously impossible, in this example, to leave exactly two numbers fixed. (Why?)

DEFINITION. A *complete permutation* is one that leaves no numbers fixed.

The problem that we now consider can be stated as follows. If a permutation of n numbers is chosen at random, what is the probability that the permutation chosen is a complete permutation? A more colorful but equivalent problem is the following. A hat-check girl has checked n hats, but they have become hopelessly scrambled. She hands back the hats at random. What is the probability that no man gets his own hat? For this problem some people's intuition would lead them to guess that for a large number of hats this probability

should be small, while others guess that it should be large. Few people guess that the probability is neither large nor small and essentially independent of the number of hats involved.

To find the desired probability, we assume that all $n!$ permutations are equally likely, and hence we need only count the number of complete permutations which there are for n elements. Let w_n be the number of such permutations. Then the desired probability is $p_n = w_n/n!$. If this procedure is carried out (see Exercise 11), the answer is found to be

$$p_n = \frac{1}{2!} - \frac{1}{3!} + \frac{1}{4!} - \cdots \pm \frac{1}{n!}$$

where the $+$ sign is chosen if n is even and the $-$ sign if n is odd. In Figure 2, these numbers are given for the first few values of n.

Number of hats	Probability p_n that no man gets his hat
2	.500000
3	.333333
4	.375000
5	.366667
6	.368056
7	.367857
8	.367882

Figure 2

It can be shown that, as the number of hats increases, the probabilities approach a number $1/e = .367879\ldots$, where the number $e = 2.718281\ldots$ is a number that plays an important role in many branches of mathematics.

EXERCISES

1. What odds should you be willing to give on a bet that at least two people in the United States Senate have the same birthday?

[*Ans.* More than 160,000:1.]

2. What is the probability that in the House of Representatives at least two men have the same birthday?

3. What odds should you be willing to give on a bet that at least two of the Presidents of the United States have had the same birthday? Would you win the bet?

[*Ans.* More than 3:1; Yes. Polk and Harding were born on Nov. 2.]

4. What odds should you be willing to give on the bet that at least two of the Presidents of the United States have died on the same day of the year? Would you win the bet?

[*Ans.* More than 2.4:1; Yes. Jefferson, Adams, and Monroe all died on July 4.]

5. Four men check their hats. Assuming that the hats are returned at random, what is the probability that *exactly* four men get their own hats? Calculate the answer for exactly 3, 2, 1, 0 men. [*Ans.* $\frac{1}{24}$; 0; $\frac{1}{4}$; $\frac{1}{3}$; $\frac{3}{8}$.]

6. A group of 50 men and their wives attend a dance. The partners for a dance are chosen by lot. What is the approximate probability that no man dances with his wife?

7. Show that the probability that, in a group of r people, *exactly* one pair has the same birthday is

$$t_r = \binom{r}{2} \frac{365 \cdot 364 \cdots (365 - r + 2)}{365^r}.$$

8. Show that $t_r = \binom{r}{2} \frac{q_r}{366 - r}$, where t_r is defined in Exercise 7, and q_r is the probability that no pair has the same birthday.

9. Using the result of Exercise 8 and the results given in Figure 1, find the probability of exactly one pair of people with the same birthday in a group of r people, for $r = 15$, 20, 25, 30, 40, and 50.

[*Ans.* .22; .32; .38; .38; .26; .12.]

10. What is the approximate probability that there has been exactly one pair of Presidents with the same birthday?

11. Let w_n be the number of complete permutations of n numbers.

(a) Show that

$$w_1 = 0, w_2 = 1, \ldots,$$

$$w_n = (n - 1)w_{n-1} + (n - 1)w_{n-2} \qquad n = 2, 3, \ldots$$

(*Hint:* Any complete permutation of n numbers can be obtained from a complete permutation of $n - 1$ numbers or from a permutation of $n - 1$ numbers that leaves one number fixed. Describe how this can be done, and show that the two terms on the right side of the equation represent the number that can be obtained from each of these methods.)

(b) Let p_n be the probability that a permutation of n numbers chosen at random is a complete permutation. From part (a) show that

$$p_1 = 0, \qquad p_2 = \frac{1}{2}$$

$$p_n = \frac{n-1}{n} p_{n-1} + \frac{1}{n} p_{n-2} \quad \text{for} \quad n = 3, 4, \ldots$$

(c) Let $v_n = p_n - p_{n-1}$ for $n = 2, 3, 4, \ldots$. From part (b), show that

$$n(p_n - p_{n-1}) = -(p_{n-1} - p_{n-2}), \qquad n = 3, \ldots,$$

and hence that

$$nv_n = -v_{n-1}, \qquad n = 3, \ldots.$$

(d) Using the fact that $p_1 = 0$, and $p_2 = \frac{1}{2}$, find v_2. From the result of part (c) find $v_3, v_4, \ldots, v_n$.

(e) Using the result of part (d), show that

$$p_n = \frac{1}{2!} - \frac{1}{3!} + \ldots \pm \frac{1}{n!}.$$

5. CONDITIONAL PROBABILITY

Suppose that we have a given $\mathfrak{U}$ and that measures have been assigned to all subsets of $\mathfrak{U}$. A statement p will have probability $\Pr[p] = m(P)$. Suppose we now receive some additional information, say that statement q is true. How does this additional information alter the probability of p?

The probability of p after the receipt of the information q is called its *conditional probability*, and it is denoted by $\Pr[p|q]$, which is read "the probability of p given q." In this section we will construct a method of finding this conditional probability in terms of the measure m.

If we know that q is true, then the original possibility set $\mathfrak{U}$ has been reduced to Q and therefore we must define our measure on the subsets of Q instead of on the subsets of $\mathfrak{U}$. Of course, every nonempty subset X of Q is a subset of $\mathfrak{U}$, and hence we know $m(X)$, its measure before q was discovered. Since q cuts down on the number of possibilities, its new measure $m'(X)$ should be larger.

The basic idea on which the definition of m' is based is that, while we know that the possibility set has been reduced to Q, we have no new information about subsets of Q. If X and Y are subsets of Q, and $m(X) = 2 \cdot m(Y)$, then we will want $m'(X) = 2 \cdot m'(Y)$. This

will be the case if the measures of subsets of Q are simply increased by a proportionality factor $m'(X) = k \cdot m(X)$, and all that remains is to determine k. Since we know that $1 = m'(Q) = k \cdot m(Q)$, we see that $k = 1/m(Q)$ and our new measure on subsets of $\mathfrak{U}$ is determined by the formula

$$m'(X) = \frac{m(X)}{m(Q)}. \tag{1}$$

How does this affect the probability of p? First of all the truth set of p has been reduced. Because all elements of $\tilde{Q}$ have been eliminated, the new truth set of p is $P \cap Q$ and therefore

$$\Pr[p|q] = m'(P \cap Q) = \frac{m(P \cap Q)}{m(Q)} = \frac{\Pr[p \wedge q]}{\Pr[q]}. \tag{2}$$

Note that if the original measure m is the equiprobable measure, then the new measure m' will also be the equiprobable measure on the set Q.

We must take care that the denominators in (1) and (2) be different from zero. Observe that $m(Q)$ will be zero if Q is the empty set, which happens only if q is self-contradictory. This is also the only case in which $\Pr[q] = 0$, and hence we make the obvious assumption that our information q is not self-contradictory.

Example 1. In an election, candidate A has a .4 chance of winning, B has .3 chance, C has .2 chance, and D has .1 chance. Just before the election C withdraws. What are now the chances of the other three candidates? Let q be the statement that C will not win, i.e., that A or B or D will win. Observe that $\Pr[q] = .8$, hence all the other probabilities are increased by a factor of $1/.8 = 1.25$. Candidate A now has .5 chance of winning, B has .375, and D has .125.

Example 2. A family is chosen at random from the set of all families having exactly two children (not twins). What is the probability that the family has two boys, if it is known that there is a boy in the family? Without any information being given, we would assign the equiprobable measure on the set $\mathfrak{U} = \{BB, BG, GB, GG\}$ where the first letter of pair indicates the sex of the younger child and the second that of the older. The information that there is a boy causes $\mathfrak{U}$ to change to $\{BB, BG, GB\}$, but the new measure is still the equi-

probable measure. Thus the conditional probability that there are two boys given that there is a boy is $\frac{1}{3}$. If on the other hand, we know that the first child is a boy, then the conditional probability is $\frac{1}{2}$.

A particularly interesting case of conditional probability is that in which $\Pr[p|q] = \Pr[p]$. Here the new information q has no effect on the probability of p, and we then say that p is *independent* of q. If in (2) we replace $\Pr[p|q]$ by $\Pr[p]$, and cross-multiply, we get

$$\Pr[p \wedge q] = \Pr[p] \cdot \Pr[q]. \tag{3}$$

On the other hand, if we express the condition that q is independent of p, we arrive at the same result. Hence the two statements are independent of each other. We can therefore say that p and q are independent if and only if case (3) holds.

Example 3. Consider three throws of an ordinary coin, where we consider the eight possibilities to be equally likely. Let p be the statement, "A head turns up on the first throw," and q be the statement, "A tail turns up on the second throw." Then $\Pr[p] = \Pr[q] = \frac{1}{2}$ and $\Pr[p \wedge q] = \frac{1}{4}$ and therefore p and q are independent statements.

While we have an intuitive notion of independence, it can happen that two statements, which may not seem to be independent, are in fact independent. For example, let r be the statement, "The same side turns up all three times." Let s be the statement "At most one head occurs." Then r and s are independent statements (see Exercise 10).

EXERCISES

1. A card is drawn at random from a pack of playing cards. What is the probability that it is a 5, given that it is between 2 and 7 inclusive?

2. There are 200 participants in a raffle. How much is one's chance of winning increased if 125 other names are eliminated?

3. A die is thrown twice. What is the probability that the sum of the faces which turn up is greater than 10, given that one of them is a 6? Given that the first throw is a 6? [*Ans.* $\frac{3}{11}$; $\frac{1}{3}$.]

4. Referring to Chapter IV, Section 3, Exercise 9, what is the probability that the man selected studies German if:

(a) He studies French?

(b) He studies French and Russian?
(c) He studies neither French nor Russian?

5. In the primary voting example of Chapter II, Section 1, assuming that the equiprobable measure has been assigned, find the probability that A wins at least two primaries, given that B drops out of the Wisconsin primary? [*Ans.* $\frac{7}{9}$.]

6. If $\Pr[\sim p] = \frac{1}{4}$ and $\Pr[q|p] = \frac{1}{2}$, what is $\Pr[p \wedge q]$? [*Ans.* $\frac{3}{8}$.]

7. A student takes a five-question true-false exam. What is the probability that he will get all answers correct if:
(a) He is only guessing?
(b) He knows that the instructor puts more true than false questions on his exams?
(c) He also knows that the instructor never puts three questions in a row with the same answer?
(d) He also knows that the first and last questions must have the opposite answer?
(e) He also knows that the answer to the second problem is "false"?

8. Three persons, A, B, and C, are placed at random in a straight line. Let r be the statement, "B is to the right of A," and let s be the statement, "C is to the right of A."
(a) What is the $\Pr[r \wedge s]$? [*Ans.* $\frac{1}{3}$.]
(b) Are r and s independent? [*Ans.* No.]

9. Let a deck of cards consist of the jacks and queens chosen from a bridge deck, and let two cards be drawn from the new deck. Find:
(a) The probability that the cards are both jacks, given that one is a jack. [*Ans.* $\frac{3}{11} = 0.27$.]
(b) The probability that the cards are both jacks, given that one is a red jack. [*Ans.* $\frac{5}{13} = 0.38$.]
(c) The probability that the cards are both jacks, given that one is the jack of hearts. [*Ans.* $\frac{3}{7} = 0.43$.]

10. Prove that statements r and s in Example 3 are independent.

11. The following example shows that r may be independent of p and q without being independent of $p \wedge q$ and $p \vee q$. We throw a coin twice. Let p be "The first toss comes out heads," q be "The second toss comes out heads," and r be "The two tosses come out the same." Compute $\Pr[r]$, $\Pr[r|p]$, $\Pr[r|q]$, $\Pr[r|p \wedge q]$, $\Pr[r|p \vee q]$. [*Ans.* $\frac{1}{2}$, $\frac{1}{2}$, $\frac{1}{2}$, 1, $\frac{1}{3}$.]

12. Prove that for any two statements p and q,

$$\Pr[p] = \Pr[p \wedge q] + \Pr[p \wedge \sim q].$$

13. Assume that p and q are independent statements. Prove that each of the following pairs of statements are independent.

(a) p and $\sim q$.
(b) $\sim q$ and p.
(c) $\sim p$ and $\sim q$.

14. Prove that for any three statements p, q, and r,

$$\Pr[p \wedge q \wedge r] = \Pr[p] \Pr[q|p] \Pr[r|p \wedge q].$$

*6. MEASURES AS AREAS

For many purposes it is convenient to represent measures by means of areas. We draw the universal set $\mathfrak{U}$ as a unit square, and to each element of $\mathfrak{U}$ we assign an area equal to its weight. By choosing these areas nonoverlapping, we will have assigned all of $\mathfrak{U}$. Then to any set we assign the total area of its elements, which will be equal to its measure.

Example 1. Let $\mathfrak{U} = \{a, b, c\}$, and assign the weights $\frac{1}{2}$, $\frac{1}{3}$, $\frac{1}{6}$ to a, b, and c, respectively. Figure 3 shows the corresponding areas, and also shows the area associated with $\{a, c\}$ (shaded in the figure). The latter area is, of course, $\frac{1}{2} + \frac{1}{6} = \frac{2}{3}$.

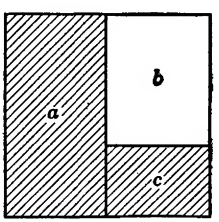


Figure 3

It is not necessary to start with the weights. We can draw in the subsets directly as long as one condition is satisfied: each set represented in our diagram has area equal to its measure.

The geometric representation helps to clarify many theoretical considerations. As an illustration we will consider the measure of the union of two sets $m(P) = .5$, $m(Q) = .3$. What is $m(P \cup Q)$? To P we assign an area of .5, and then we want to add Q to our diagram. The set Q is assigned an area of .3, but how much is this area to overlap P? If (a) there is no overlap, then the total area is .8; while (b) if Q is inside P, then the total area is .5; finally (c) if the overlap is, say, equal to .2, then the total area is .6. (These three possibilities are shown in Figure 4.) Since, in each case, the overlap represents $P \cap Q$, we have no choice: we must make the area of the overlap equal to $m(P \cap Q)$. It is easy to read the formula $m(P \cup Q) =$

$m(P) + m(Q) - m(P \cap Q)$ from Figure 4(c). Figure 4(a) shows the case where P and Q are disjoint.

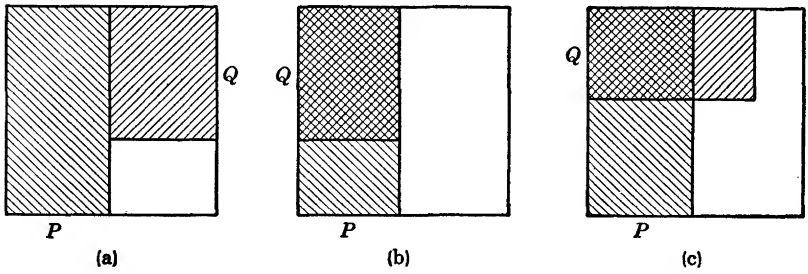


Figure 4

If we have just one subset of $\mathfrak{U}$, say P, we can always represent it by a vertical strip, i.e., a rectangle with height equal to one. But if Q is added, we cannot always represent it by a complete horizontal strip; as can be seen in Figure 4. Let us consider the special case, shown in Figure 5, where such representation is possible. The set P has base equal to $m(P)$ and height one. The set Q has base one and height $m(Q)$. The intersection, $P \cap Q$, has area $m(P) \cdot m(Q)$. That means that $m(P \cap Q) = m(P) \cdot m(Q)$ which is the special case of independence.

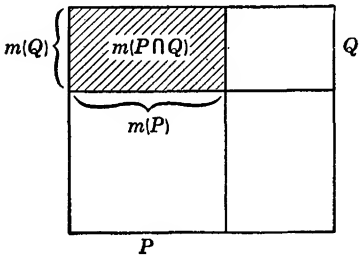


Figure 5

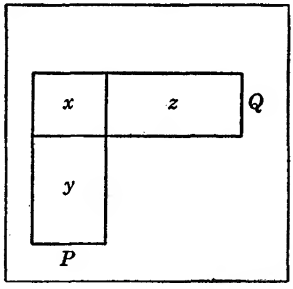


Figure 6

To represent the probabilities of statements, we must represent the measures of their truth sets. Hence the above method is applicable to the representation of probabilities as well. It can be used also to represent conditional probabilities: $\Pr[p|q] = \Pr[p \wedge q]/\Pr[q] =$

$m(P \cap Q)/m(Q)$. Hence the conditional probability is represented by the ratio of the common area of P and Q to the total area of Q. In Figure 6 let $m(P) = x + y$, $m(Q) = x + z$, and $m(P \cap Q) = x$; then $\Pr[p|q] = x/(x + z)$ and $\Pr[q|p] = x/(x + y)$.

As an application of this geometric representation we will develop Bayes' theorem. The simplest case of this theorem is where we know the probability of p occurring, and later we get additional evidence q whose relation to p is known. The question is how the new information changes the probability of p. More specifically, we know $\Pr[p]$, $\Pr[q|p]$, and $\Pr[q|\sim p]$ to start with. What we want to know is $\Pr[p|q]$.

Example 2. Suppose we have two urns. The first contains two black balls and one white ball, while the second contains two white balls and one black ball. We select an urn according to a random device, which makes the probability of choosing the first urn $\frac{3}{4}$, and then draw a ball. If a black ball is drawn, what is the probability that we drew from the first urn? Let p state that we draw from the first urn, and q that we draw a black ball. $\Pr[p] = \frac{3}{4}$, $\Pr[q|p] = \frac{2}{3}$, $\Pr[q|\sim p] = \frac{1}{3}$. We are to find $\Pr[p|q]$. These numbers are represented in Figure 7. Observe that $x_1 + y_1 = \frac{3}{4}$, $x_2 + y_2 = \frac{1}{4}$, $x_1/(x_1 + y_1) = \frac{2}{3}$, and $x_2/(x_2 + y_2) = \frac{1}{3}$. Therefore, $x_1 = (\frac{3}{4})(\frac{2}{3}) = \frac{1}{2}$, and $x_2 = (\frac{1}{3})(\frac{1}{4}) = \frac{1}{12}$. Finally $\Pr[p|q] = x_1/(x_1 + x_2) = \frac{6}{7}$.

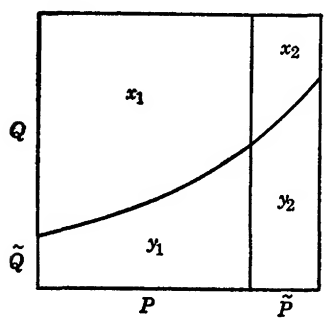


Figure 7

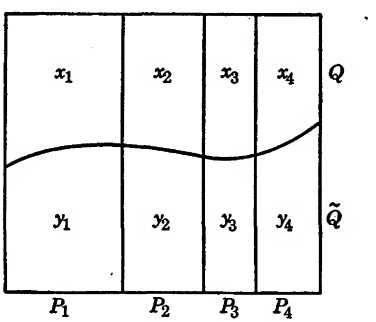


Figure 8

Here we started with only two alternatives, p and $\sim p$. But the theorem to be developed is applicable to any number of alternatives. We will work it out for four alternatives. The statements p_1, p_2, p_3,

and p_4 are said to form a complete set of alternatives if one of them must be true and no more than one can be true. (See Chapter I, Section 8.) In this case their truth sets are disjoint, and their union is $\mathfrak{U}$. Let q be some other statement. This situation is represented in Figure 8.

As initial data we are given the probabilities of the four alternatives, and we know the conditional probability of q relative to each alternative.

$$\begin{aligned} \Pr[p_1] &= x_1 + y_1 & \Pr[q|p_1] &= x_1/(x_1 + y_1) \\ \Pr[p_2] &= x_2 + y_2 & \Pr[q|p_2] &= x_2/(x_2 + y_2) \\ \Pr[p_3] &= x_3 + y_3 & \Pr[q|p_3] &= x_3/(x_3 + y_3) \\ \Pr[p_4] &= x_4 + y_4 & \Pr[q|p_4] &= x_4/(x_4 + y_4) \end{aligned}$$

Thus we are given the areas of the four vertical strips and know in each case what fraction the upper portion of the strip is, so that we have the area of the four upper portions: By multiplying the two probabilities on each line we obtain $x_1 = \Pr[p_1] \cdot \Pr[q|p_1]$, $x_2 = \Pr[p_2] \cdot \Pr[q|p_2]$, etc. We are interested in finding the probability of one of the four alternatives, say p_2, given that q has taken place. In other words we want to know what fraction x_2 is of the area of Q. The formula

$$\Pr[p_2|q] = \frac{\Pr[p_2] \cdot \Pr[q|p_2]}{\Pr[p_1] \cdot \Pr[q|p_1] + \Pr[p_2] \cdot \Pr[q|p_2] + \Pr[p_3] \cdot \Pr[q|p_3] + \Pr[p_4] \cdot \Pr[q|p_4]}$$

gives the desired probability, as can be checked. Similar formulas apply for the other alternatives, and the formula generalizes in an obvious way to any number of alternatives. In its most general form it is called Bayes' theorem.

Example 3. Suppose that a freshman must choose among mathematics, physics, chemistry, and astronomy as his science course. On the basis of the interest he expressed, his adviser assigns probabilities of .4, .3, .2, and .1 to his choosing each of the four courses, respectively. His adviser does not hear which course he actually chose, but at the end of the term the adviser hears that he received A in the course chosen. On the basis of the difficulties of these courses the adviser estimates the probability of the student getting an A in mathematics to be .1, in physics .2, in chemistry .3, and in astronomy .9. How can the adviser revise his original estimates as to the prob-

abilities of his taking the various courses? Using Bayes' theorem we get

$$\Pr[\text{He took Math}|\text{He got an A}] = \frac{(.4)(.1)}{(.4)(.1) + (.3)(.2) + (.2)(.3) + (.1)(.9)} = \frac{4}{25} = .16.$$

Similar computations assign probabilities of .24, .24, and .36 to the other three courses. Thus the new information, that he received an A, had little effect on the probability of his having taken physics or chemistry, but it has made it much less likely that he took mathematics, and much more likely that he took astronomy.

It is important to note that knowing the conditional probabilities of q relative to the alternatives is not enough. Unless we also know the probabilities of the alternatives at the start, we cannot apply Bayes' theorem. However, in some situations it is reasonable to assume that the alternatives are equally probable at the start. In this case the factors $\Pr[p_1], \ldots, \Pr[p_4]$ cancel from our basic formula, and we get the special form of the theorem:

If $\Pr[p_1] = \Pr[p_2] = \Pr[p_3] = \Pr[p_4]$, then

$$\Pr[p_2|q] = \frac{\Pr[q|p_2]}{\Pr[q|p_1] + \Pr[q|p_2] + \Pr[q|p_3] + \Pr[q|p_4]}.$$

Example 4. In a sociological experiment the subjects are handed four sealed envelopes, each containing a problem. They are told to open one envelope and try to solve the problem in ten minutes. From past experience, the experimenter knows that the probability of their being able to solve the hardest problem is .1. With the other problems they have probabilities of .3, .5, and .8. Assume the group succeeds within the allotted time. What is the probability that they selected the hardest problem? Since they have no way of knowing which problem is in which envelope, they choose at random, and we assign equal probabilities to the selection of the various problems. Hence the above simple formula applies. The probability of their having selected the hardest problem is $.1/(.1 + .3 + .5 + .8) = \frac{1}{17}$.

Example 5. Suppose that we can play any one of three slot machines, which usually pay off only with probability .1. Suppose also

that a friend has told us that one of the machines is out of order and pays off with probability .4. We select one machine at random, and play. How does the outcome affect the probability that we have selected the profitable machine? If we win, then the probability that we have selected the right machine is $.4/(.4 + .1 + .1) = \frac{2}{3}$. If we lose, the probability is $.6/(.6 + .9 + .9) = \frac{1}{4}$. Since we started with probability $\frac{1}{3}$, we see that a win gives us much more information than a loss. This is reasonable, since a loss is quite likely even on the profitable machine.

EXERCISES

1. Represent the sets P and Q as areas of a square, given the following information:

(a) $m(P) = m(Q) = .5$, $m(P \cap Q) = .3$.
(b) $m(P) = .5$, $m(Q) = .4$, $m(P \cup Q) = .8$.
(c) $m(P) = .6$, $m(Q) = .4$, P and Q are disjoint.

2. Given the following information about statements p and q, represent their truth sets P and Q as areas:

(a) $\Pr[p] = \Pr[q] = .4$, $\Pr[p \wedge q] = .2$.
(b) $\Pr[p] = \Pr[q] = .5$, $\Pr[p \vee q] = .75$.
(c) $\Pr[p] = \Pr[q] = \Pr[p \wedge q] = .3$.
(d) $\Pr[p] = \Pr[q] = .4$, $\Pr[q|p] = .5$.

3. During the month of May the probability of a rainy day is .2. The Dodgers win on a clear day with probability .7, but on a rainy day only with probability .4. If we know that they won a certain game in May, what is the probability that it rained on that day? [*Ans.* $\frac{1}{8}$.]

4. Construct a diagram to represent the truth sets of various statements occurring in the previous exercise.

5. On a multiple-choice exam there are four possible answers for each question. Therefore, if a student knows the right answer, he has probability one of choosing correctly; if he is guessing, he has probability $\frac{1}{4}$ of choosing correctly. Let us further assume that a good student will know 90 per cent of the answers, a poor student only 50 per cent. If a good student has the right answer, what is the probability that he was only guessing? Answer the same question about a poor student, if the poor student has the right answer. [*Ans.* $\frac{1}{37}$, $\frac{1}{5}$.]

6. Three economic theories are proposed at a given time, which appear to be equally likely on the basis of existing evidence. The state of the American economy is observed the following year, and it turns out that its actual de-

velopment had probability .6 of happening according to the first theory; and probabilities .4 and .2 according to the others. How does this modify the probabilities of correctness of the three theories?

7. Let p_1, p_2, p_3, and p_4 be a set of equally likely alternatives. Let $\Pr[q|p_1] = a$, $\Pr[q|p_2] = b$, $\Pr[q|p_3] = c$, $\Pr[q|p_4] = d$. Show that, if $a + b + c + d = 1$, then the revised probabilities of the alternatives relative to q are a, b, c, and d, respectively.

8. In poker, Smith holds a very strong hand and bets a considerable amount. The probability that his opponent, Jones, has a better hand is .05. With a better hand Jones would raise the bet with probability .9, but with a poorer hand Jones would raise only with probability .2. Suppose that Jones raises, what is the new probability that he has a winning hand? [*Ans.* $\frac{9}{47}$.]

9. A rat is allowed to choose one of five mazes at random. If we know that the probabilities of his getting through the various mazes in three minutes are .6, .3, .2, .1, .1, and we find that the rat escapes in three minutes, how probable is it that he chose the first maze? The second maze? [*Ans.* $\frac{6}{13}$, $\frac{3}{13}$.]

7. FINITE STOCHASTIC PROCESSES

We consider here a very general situation which we will specialize in later sections. We deal with a sequence of experiments where the outcome on each particular experiment depends on some chance element. Any such sequence is called a *stochastic process*. (The Greek word "stochos" means "guess"). We shall assume a finite number of experiments and a finite number of possibilities for each experiment. We assume that, if all the outcomes of the experiments which precede a given experiment were known, then both the possibilities for this experiment and the probability that any particular possibility will occur would be known. We wish to make predictions about the process as a whole. For example, in the case of repeated throws of an ordinary coin we would assume that on any particular experiment we have two outcomes, and the probabilities for each of these outcomes is one-half regardless of any other outcomes. We might be interested, however, in the probabilities of statements of the form, "More than two-thirds of the throws result in heads," or "The number of heads and tails which occur is the same," etc. These are questions which can be answered only when a probability measure has been assigned to the process as a whole. In this section we show how probability measure can be assigned, using the given information. In the case

of coin tossing, the probabilities (hence also the possibilities) on any given experiment do not depend upon the previous results. We will not make any such restriction here since the assumption is not true in general.

We shall show how the probability measure is constructed for a particular example, and the procedure in the general case is similar.

We assume that we have a sequence of three experiments, the possibilities for which are indicated in Figure 9. The set of all possible outcomes which might occur on any of the experiments is represented by the set $\{a, b, c, d, e, f\}$. Note that if we know that outcome b occurred on the first experiment, then we know that the possibilities on experiment two are $\{a, e, d\}$. Similarly if we know that b occurred on the first experiment and a on the second, then the only possibilities for the third are $\{c, f\}$. We denote by p_a the probability that the first experiment results in outcome a, and by p_b the probability that outcome b occurs in the first experiment. We denote by $p_{b,d}$ the probability that outcome d occurs on the second experiment, which is the probability computed on the assumption that outcome b occurred on the first experiment. Similarly for $p_{b,a}$, $p_{b,e}$, $p_{a,a}$, $p_{a,c}$. We denote by $p_{bd,c}$ the probability that outcome c occurs on the third experiment, the latter probability being computed on the assumption that outcome b occurred on the first experiment and d on the second. Similarly for $p_{ba,c}$, $p_{ba,f}$, etc. We have assumed that these numbers are given and the fact that they are probabilities assigned to possible outcomes would mean that they are positive and that

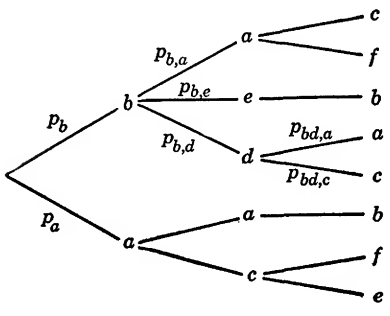


Figure 9

$$p_a + p_b = 1, \quad p_{b,a} + p_{b,e} + p_{b,d} = 1, \quad \text{and} \quad p_{bd,a} + p_{bd,c} = 1, \text{ etc.}$$

It is convenient to associate each probability with the branch of the tree that connects the branch point representing the predicted outcome. We have done this in Figure 9 for several branches. The sum of the numbers assigned to branches from a particular branch point is one, e.g., $p_{b,a} + p_{b,e} + p_{b,d} = 1$.

A possibility for the sequence of three experiments is indicated by a path through the tree. We define now a probability measure on the set of all paths. We call this a *tree measure*. To the path corresponding to outcome b on the first experiment, d on the second, and c on the third, we assign the weight $p_b \cdot p_{b,d} \cdot p_{bd,c}$. That is the *product* of the probabilities associated with each branch along the path being considered. We find the probability for each path through the tree.

Before showing the reason for this choice we must first show that it determines a probability measure, in other words that the weights are positive and the sum of the weights is one. The weights are products of positive numbers and hence positive. To see that their sum is one we first find the sum of the weights of all paths corresponding to a particular outcome, say b, on the first experiment and a particular outcome, say d, on the second. We have

$$p_b \cdot p_{b,d} \cdot p_{bd,a} + p_b \cdot p_{b,d} \cdot p_{bd,c} = p_b \cdot p_{b,d}[p_{bd,a} + p_{bd,c}] = p_b \cdot p_{b,d}.$$

For any other first two outcomes we would obtain a similar result. For example, the sum of the weights assigned to paths corresponding to outcome a on the first experiment and c on the second is $p_a \cdot p_{a,c}$. Notice that when we have verified that we have a probability measure, this will be the probability that the first outcome results in a and the second experiment results in c.

Next we find the sum of the weights assigned to all the paths corresponding to the cases where the outcome of the first experiment is b. We find this by adding the sums corresponding to the different possibilities for the second experiment. But by our preceding calculation this is

$$p_b \cdot p_{b,a} + p_b \cdot p_{b,e} + p_b \cdot p_{b,d} = p_b[p_{b,a} + p_{b,e} + p_{b,d}] = p_b.$$

Similarly the sum of the weights assigned to paths corresponding to the outcome a on the first experiment is p_a. Thus the sum of all weights is $p_a + p_b = 1$. Therefore we do have a probability measure. Note that we have also shown that the probability that the outcome of the first experiment is a has been assigned probability p_a in agreement with our given probability.

To see the complete connection of our new measure with the given probabilities, let $X_j = z$ be the statement "The outcome of the jth experiment was z." Then the statement $[X_1 = b \wedge X_2 = d \wedge X_3 = c]$ is a compound statement that has been assigned probability

$p_b \cdot p_{b,d} \cdot p_{bd,c}$. The statement $[X_1 = b \wedge X_2 = d]$ we have noted has been assigned probability $p_b \cdot p_{b,d}$ and the statement $[X_1 = b]$ has been assigned probability p_b. Thus

$$\Pr[X_3 = c | X_2 = d \wedge X_1 = b] = \frac{p_b \cdot p_{b,d} \cdot p_{bd,c}}{p_b \cdot p_{b,d}} = p_{bd,c}$$

$$\Pr[X_2 = d | X_1 = b] = \frac{p_b \cdot p_{b,d}}{p_b} = p_{b,d}.$$

Thus we see that our probabilities, computed under the assumption that previous results were known, become the corresponding conditional probabilities when computed with respect to the tree measure. It can be shown that the tree measure which we have assigned is the only one which will lead to this agreement. We can now find the probability of any statement concerning the stochastic process from our tree measure.

Example 1. Suppose that we have two urns. Urn 1 contains two black balls and three white balls. Urn 2 contains two black balls and one white ball. An urn is chosen at random and a ball chosen from this urn at random. What is the probability that a white ball is chosen? A hasty answer might be $\frac{1}{2}$, since there are an equal number of black and white balls involved and everything is done at random. However, it is hasty answers like this (which is wrong) which show the need for a more careful analysis.

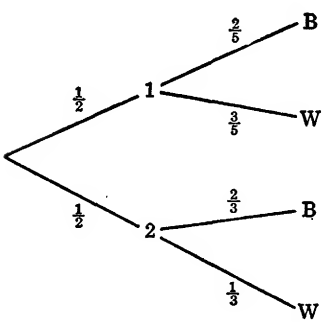


Figure 10

We are considering two experiments. The first consists in choosing the urn and the second in choosing the ball. There are two possibilities for the first experiment, and we assign $p_1 = p_2 = \frac{1}{2}$ for the probabilities of choosing the first and the second urn, respectively. We then assign $p_{1,W} = \frac{3}{5}$ for the probability that a white ball is chosen, under the assumption that urn 1 is chosen. Similarly we assign $p_{1,B} = \frac{2}{5}$, $p_{2,W} = \frac{1}{3}$, $p_{2,B} = \frac{2}{3}$. We indicate these probabilities on the possibility tree in Figure 10. The probability that a white ball is drawn is then found from the tree measure as the sum of the weights assigned to paths which lead to a choice of a white ball. This is $\frac{1}{2} \cdot \frac{3}{5} + \frac{1}{2} \cdot \frac{1}{3} = \frac{7}{15}$.

Example 2. Suppose that a man leaves a bar which is on a corner which he knows to be one block from his home. He is unable to remember which street leads to his home. He proceeds to try each of the streets at random without ever choosing the same street twice

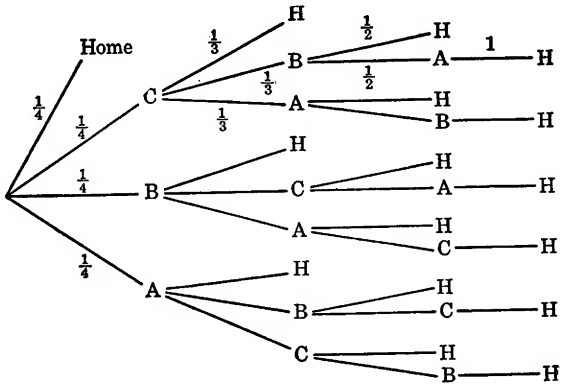


Figure 11

until he goes on the one which leads to his home. What possibilities are there for his trip home, and what is the probability for each of these possible trips? We label the streets A, B, C, and Home. The possibilities together with typical probabilities are given in Figure 11. The probability for any particular trip, or path, is found by taking the product of the branch probabilities.

Example 3. Assume that you are presented with two slot machines, A and B. Each machine pays the same fixed amount when it pays off. Machine A pays off each time with probability $\frac{1}{2}$, and machine B with probability $\frac{1}{4}$. You are not told which machine is A. Suppose that you choose a machine at random and win. What is the probability that you chose machine A? We first construct the tree (Figure 12) to show the possibilities and assign branch probabilities to determine a tree measure. Let p be the statement, "Machine

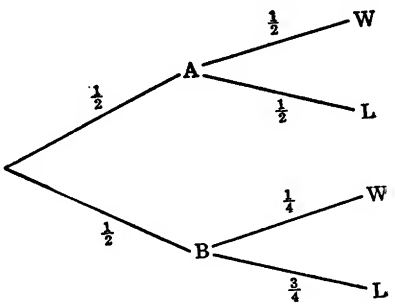


Figure 12

A was chosen," and q be the statement, "The machine chosen paid off." Then we are asked for $\Pr[p|q] = \dfrac{\Pr[p \wedge q]}{\Pr[q]}$. The truth set of the statement $p \wedge q$ consists of a single path which has been assigned weight $\frac{1}{4}$. The truth set of the statement q consists of two paths, and the sum of the weights of these paths is $\frac{1}{2}\cdot\frac{1}{2} + \frac{1}{2}\cdot\frac{1}{4} = \frac{3}{8}$. Thus $\Pr[p|q] = \frac{2}{3}$. Thus if we win it is more likely that we have machine A than B and suggests that next time we should play the same machine. If we lose, however, it is more likely that we have machine B than A, and hence we would switch machines before the next play. (See Exercise 9.)

EXERCISES

1. The fractions of Republicans, Democrats, and Independent voters in cities A and B are

City A: .30 Republican, .40 Democratic, .30 Independent;
City B: .40 Republican, .50 Democratic, .10 Independent.

A city is chosen at random and two voters are chosen successively and at random from the voters of this city. Construct a tree measure and find the probability that two Democrats are chosen. Find the probability that the second voter chosen is an Independent voter. [*Ans.* .205; .2.]

2. A coin is thrown. If a head turns up a die is rolled. If a tail turns up the coin is thrown again. Construct a tree measure to represent the two experiments and find the probability that the die is thrown and a six turns up.

3. A man wins a certain tournament if he can win two consecutive games out of three played alternately with two opponents A and B. A is a better player than B. The probability of winning a game when B is the opponent is $\frac{2}{3}$. The probability of winning a game when A is his opponent is only $\frac{1}{3}$. Construct a tree measure for the possibilities for three games, assuming that he plays alternately but plays A first. Do the same assuming that he plays B first. In each case find the probability that he will win two consecutive games. Is it better to play two games against the strong player or against the weaker player? [*Ans.* $\frac{10}{27}$; $\frac{8}{27}$; better to play strong player twice.]

4. Construct a tree measure to represent the possibilities for four throws of an ordinary coin. Assume that the probability of a head on any toss is $\frac{1}{2}$ regardless of any information about other throws.

5. A student claims to be able to distinguish beer from ale. He is given a series of three tests. In each test he is given two cans of beer and one of

ale and asked to pick out the ale. If he gets two or more correct we will admit his claim. Draw a tree to represent the possibilities (either right or wrong) for his answers. Construct the tree measure which would correspond to guessing and find the probability that his claim will be established if he guesses on every trial.

6. A box contains three defective light bulbs and seven good ones. Construct a tree to show the possibilities if three consecutive bulbs are drawn at random from the box (they are not replaced after being drawn). Assign a tree measure and find the probability that at least one good bulb is drawn out. Find the probability that all three are good if the first bulb is good.
[*Ans.* $\frac{11}{120}$; $\frac{5}{12}$.]

7. In Example 2 above, find the probability that the man reaches home after trying at most one wrong street.

8. In Example 3, find the probability that machine A was chosen, given that the player lost.

9. In Example 3, assume that the player makes two plays. Find the probability that he wins at least once under the assumption:

(a) That he plays the same machine twice. [*Ans.* $\frac{18}{32}$.]

(b) That he plays the same machine the second time if and only if he won the first time. [*Ans.* $\frac{20}{32}$.]

10. A chess player plays three successive games of chess. His psychological makeup is such that the probability of his winning a given game is $(\frac{1}{2})^{k+1}$, where k is the number of games he has won so far. (For instance, the probability of his winning the first game is $\frac{1}{2}$, the probability of his winning the second game *if he has already won the first game* is $\frac{1}{4}$, etc.) What is the probability that he will win at least two of the three games?

11. Before a political convention, a political expert has assigned the following probabilities. The probability that the President will be willing to run again is $\frac{1}{2}$. If he is willing to run, he and his Vice President are sure to be nominated and have probability $\frac{3}{5}$ of being elected again. If the President does not run, the present Vice President has probability $\frac{1}{10}$ of being nominated, and any other presidential candidate has probability $\frac{1}{2}$ of being elected. What is the probability that the present Vice President will be re-elected?
[*Ans.* $\frac{13}{40}$.]

12, 13. Work Exercises 3 and 8 of Section 6, using the method illustrated in Example 3 of this section.

14. There are two urns, A and B. Urn A contains one black and one red ball. Urn B contains two black and three red balls. A ball is chosen at random from urn A and put into urn B. A ball is then drawn at random from urn B.

(a) What is the probability that both balls drawn are of the same color? [*Ans.* $\frac{7}{12}$.]
(b) What is the probability that the first ball drawn was red, given that the second ball drawn was black? [*Ans.* $\frac{2}{5}$.]

15. Assume that in the World Series each team has probability one-half of winning each game, independently of the outcomes of any other game. Assign a tree measure. (See Chapter 1, Section 6 for the tree.) Find the probability that the series ends in 4, 5, 6, and 7 games, respectively.

16. Assume that in the World Series one team is stronger than the other and has probability $\frac{2}{3}$ for winning each of the games. Assign a tree measure and find the following probabilities.
(a) The probability that the stronger team wins in 4, 5, 6, and 7 games, respectively.
(b) The probability that the weaker team wins in 4, 5, 6, and 7 games, respectively.
(c) The probability that the series ends in 4, 5, 6, and 7 games, respectively. [*Ans.* .21; .30; .27; .22.]
(d) The probability that the strong team wins the series. [*Ans.* .83.]

17. In the World Series from 1905 to 1955, excluding the nine-game series, there have been 10 four-game series, 13 five-game series, 12 six-game series, and 13 seven-game series. Add the results from 1955 to date to these and use these past records to estimate the probability that a series will last 4, 5, 6, or 7 games. Compare your answers with those obtained theoretically in Exercises 15 and 16(c). Which assumption about the World Series play seems to fit the data better?

8. INDEPENDENT TRIALS WITH TWO OUTCOMES

In the preceding section we developed a way to determine a probability measure for any sequence of chance experiments where there are only a finite number of possibilities for each experiment. While this provides the framework for the general study of stochastic processes, it is too general to be studied in complete detail. Therefore, in probability theory we look for simplifying assumptions which will make our probability measure easier to work with. It is desired also that these assumptions be such as to apply to a variety of experiments which would occur in practice. In this book we shall limit ourselves to the study of two types of processes. The first, the independent trials process, will be considered in the present section. This process was the first one to be studied extensively in probability theory. The

second, the Markov chain process, is a process that is finding increasing application, particularly in the social and biological sciences, and will be considered in Section 13.

A process of independent trials applies to the following situation. Assume that there is a sequence of chance experiments, each of which consists of a repetition of a single experiment, carried out in such a way that the results of any one experiment in no way affect the results in any other experiment. We label the possible outcome of a single experiment by $a_1, \ldots, a_r$. We assume that we are also given probabilities $p_1, \ldots, p_r$ for each of these outcomes occurring on any single experiment, the probabilities being independent of previous results. The tree representing the possibilities for the sequence of experiments will have the same outcomes from each branch point, and the branch probabilities will be assigned by assigning probability p_j to any branch leading to outcome a_j. The tree measure determined in this way is the measure of an *independent trials process.* In this section we shall consider the important case of two outcomes for each experiment. The more general case is studied in Section 11.

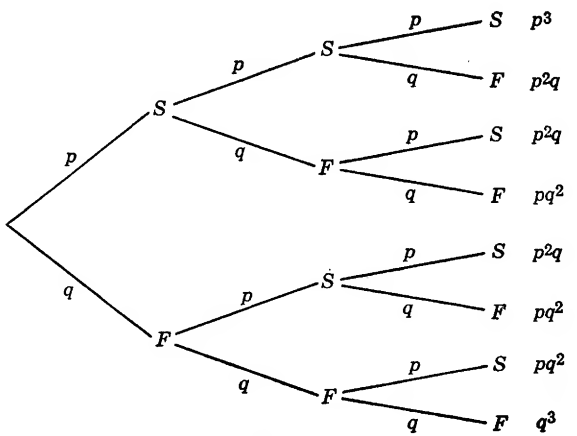


Figure 13

In the case of two outcomes we arbitrarily label one outcome "success" and the other "failure." For example, in repeated throws of a coin we might call heads success, and tails failure. We assume there is given a probability p for success and a probability $q = 1 - p$

for failure. The tree measure for a sequence of three such experiments is shown in Figure 13. The weights assigned to each path are indicated at the end of the path.

The question which we now ask is the following. Given an independent trials process with two outcomes, what is the probability of *exactly* x successes in n experiments. We denote this probability by $f(n,x;p)$ to indicate that it depends upon n, x, and p.

Assume that we had a tree for this general situation, similar to the tree in Figure 13 for three experiments, with the branch points labeled S for success and F for failure. Then the truth set of the statement, "Exactly x successes occur," consists of all paths which go through x branch points labeled S and $n - x$ labeled F. To find the probability of this statement we must add the weights for all such paths. We are helped first by the fact that our tree measure assigns the same weight to any such path, namely $p^x q^{n-x}$. The reason for this is that every branch leading to an S is assigned probability p, and every branch leading to F is assigned probability q, and in the product there will be x p's and $n - x$ q's. To find the desired probability we need only find the number of paths in the truth set of the statement, "Exactly x successes occur." To each such path we make correspond an ordered partition of the integers from 1 to n which has two cells, x elements in the first and $n - x$ in the second. We do this by putting the numbers of the experiments on which success occurred in the first cell and those for which failure occurred in the second cell. Since there are $\binom{n}{x}$ such partitions there are also this number of paths in the truth set of the statement considered. Thus we have proved:

In an independent trials process with two outcomes the probability of exactly x successes in n experiments is given by

$$f(n,x;p) = \binom{n}{x} p^x q^{n-x}.$$

Example 1. Consider n throws of an ordinary coin. We label heads "success" and tails "failure," and we assume that the probability is $\frac{1}{2}$ for heads on any one throw independently of the outcome of any other throw. Then the probability that exactly x heads will turn up is

$$f(n,x;\tfrac{1}{2}) = \binom{n}{x}\left(\frac{1}{2}\right)^n.$$

For example, in 100 throws the probability that exactly 50 heads will turn up is $f(100,50;\frac{1}{2}) = \binom{100}{50}(\frac{1}{2})^{100}$ which is approximately .08. Thus we see that it is quite unlikely that exactly one-half of the tosses will result in heads. On the other hand, suppose that we ask for the probability that nearly one-half of the tosses will be heads. To be more precise, let us ask for the probability that the number of heads which occur does not deviate by more than 10 from 50. To find this we must add $f(100,x;\frac{1}{2})$ for $x = 40, 41, \ldots, 60$. If this is done, we obtain a probability of approximately .96. Thus, while it is unlikely that exactly 50 heads will occur, it is very likely that the number of heads which occur will not deviate from 50 by more than 10.

Example 2. Assume that we have a machine which, on the basis of data given, is to predict the outcome of an election as either a Republican victory or a Democratic victory. If two identical machines are given the same data, they should predict the same result. We assume, however, that any such machine has a certain probability q of reversing the prediction that it would ordinarily make, because of a mechanical or electrical failure. To improve the accuracy of our prediction we give the same data to r identical machines, and choose the answer which the majority of the machines give. To avoid ties we assume that r is odd. Let us see how this decreases the probability of an error due to a faulty machine.

Consider r experiments, where the jth experiment results is success if the jth machine produces the prediction which it would make when operating without any failure of parts. The probability of success is then $p = 1 - q$. The majority decision will agree with that of a perfectly operating machine if we have more than $r/2$ successes. Suppose, for example, that we have five machines, each of which has a probability of .1 of reversing the prediction because of a parts failure. Then the probability for success is .9, and the probability that the majority decision will be the desired one is

$$f(5,3;0.9) + f(5,4;0.9) + f(5,5;0.9)$$

which is found to be approximately .991 (see Exercise 3).

Thus the above procedure decreases the probability of error due to machine failure from .1 in the case of one machine to .009 for the case of five machines.

EXERCISES

1. Compute for $n = 4$, $n = 8$, $n = 12$, and $n = 16$ the probability of obtaining exactly $\frac{1}{2}$ heads when an ordinary coin is thrown.

[*Ans.* .375; .273; .226; .196.]

2. Compute for $n = 4$, $n = 8$, $n = 12$, and $n = 16$ the probability that the fraction of heads deviates from $\frac{1}{2}$ by less than $\frac{1}{5}$.

[*Ans.* .375; .711; .854; .923.]

3. Verify that the probability .991 given in Example 2 is correct.

4. Assume that Peter and Paul match pennies four times. (In matching pennies, Peter wins a penny with probability $\frac{1}{2}$, and Paul wins a penny with probability $\frac{1}{2}$.) What is the probability that Peter wins more than Paul? Answer the same for five throws. For the case of 12,917 throws.

[*Ans.* $\frac{5}{16}$; $\frac{1}{2}$; $\frac{1}{2}$.]

5. If an ordinary die is thrown four times, what is the probability that exactly two 6's will occur?

6. In a ten-question true-false exam, what is the probability of getting 70 per cent or better by guessing? [*Ans.* $\frac{11}{64}$.]

7. Assume that, every time a batter comes to bat, he has probability .3 for getting a hit. Assuming that his hits form an independent trials process and that the batter comes to bat four times, what fraction of the games would he expect to get at least two hits? At least three hits? Four hits?

[*Ans.* .348; .084; .008.]

8. A coin is to be thrown eight times. What is the most probable number of heads that will occur? What is the number having the highest probability, given that the first four throws resulted in heads?

9. A small factory has ten workers. The workers eat their lunch at one of two diners, and they are just as likely to eat in one as in the other. If the proprietors want to be more than .95 sure of having enough seats, how many seats must each of the diners have? [*Ans.* Eight seats.]

10. Suppose that five people are chosen at random and asked if they favor a certain proposal. If only 30 per cent of the people favor the proposal, what is the probability that a majority of the five people chosen will favor the proposal?

11. In Example 2, if the probability for a machine reversing its answer due to a parts failure is .2, how many machines would have to be used to make the probability greater than .89 that the answer obtained would be that which a machine with no failure would give? [*Ans.* Three machines.]

12. Assume that it is estimated that a torpedo will hit a ship with probability $\frac{1}{3}$. How many torpedoes must be fired if it is desired that the probability for at least one hit should be greater than .9?

13. A student estimates that, if he takes four courses, he has probability .8 of passing each course. If he takes five courses he has probability .7 of passing each course, and if he takes six courses he has probability .5 for passing each course. His only goal is to pass at least four courses. How many courses should he take for the best chance of achieving his goal? [*Ans.* 5.]

*9. A PROBLEM OF DECISION

In the preceding sections we have dealt with the problem of calculating the probability of certain statements based on the assumption of a given probability measure. In a statistics problem, one is often called upon to make a decision in a case where the decision would be relatively easy to make if we could assign probabilities to certain statements, but we do not know how to assign these probabilities. For example, if a vaccine for a certain disease is proposed, we may be called upon to decide whether or not the vaccine should be used. We may decide that we could make the decision if we could compare the probability that a person vaccinated will get the disease with the probability that a person not vaccinated will get the disease. Statistical theory develops methods to obtain from experiments some information which will aid in estimating these probabilities, or will otherwise help in making the required decision. We shall illustrate a typical procedure.

Smith claims that he has the ability to distinguish ale from beer and has bet Jones a dollar to that effect. Now Smith does not mean that he can distinguish beer from ale with 100 per cent accuracy, but rather that he believes that he can distinguish them a proportion of the time which is significantly greater than $\frac{1}{2}$.

Assume that it is possible to assign a number p which represents the probability that Smith can pick out the ale from a pair of glasses, one containing ale and one beer. We identify $p = \frac{1}{2}$ with his having no ability, $p > \frac{1}{2}$ with his having some ability, and $p < \frac{1}{2}$ with his being able to distinguish, but having the wrong idea which is the ale. If we knew the value of p, we would award the dollar to Jones if p were $\leq \frac{1}{2}$, and to Smith if p were $> \frac{1}{2}$. As it stands, we have no knowledge of p and thus cannot make a decision. We perform an experiment and make a decision as follows.

Smith is given a pair of glasses, one containing ale and the other beer, and is asked to identify which is the ale. This procedure is repeated ten times, and the number of correct identifications is noted. If the number correct is at least eight, we award the dollar to Smith, and, if it is less than eight, we award the dollar to Jones.

We now have a definite procedure and shall examine this procedure both from Jones's and Smith's points of view. We can make two kinds of errors. We may award the dollar to Smith when in fact the appropriate value of p is $\leq \frac{1}{2}$, or we may award the dollar to Jones when the appropriate value for p is $> \frac{1}{2}$. There is no way that these errors can be completely avoided. We hope that our procedure is such that each of the bettors will be convinced that, if he is right, he will very likely win the bet.

Jones believes that the true value of p is $\frac{1}{2}$. We shall calculate the probability of Jones winning the bet if this is indeed true. We assume that the individual tests are independent of each other and all have the same probability $\frac{1}{2}$ for success. (This assumption will be unreasonable if the glasses are too large.) We have then an independent trials process with $p = \frac{1}{2}$ to describe the entire experiment. The probability that Jones will win the bet is the probability that Smith gets fewer than eight correct. From the table in Figure 14 we com-

x \ p	0.1	0.25	0.50	0.75	0.90
0	.349	.056	.001	.000	.000
1	.387	.188	.010	.000	.000
2	.194	.282	.044	.000	.000
3	.057	.250	.117	.003	.000
4	.011	.146	.205	.016	.000
5	.001	.058	.246	.058	.001
6	.000	.016	.205	.146	.011
7	.000	.003	.117	.250	.057
8	.000	.000	.044	.282	.194
9	.000	.000	.010	.188	.387
10	.000	.000	.001	.056	.349

Table of Values of $f(10,x;p)$

Figure 14

pute that this probability is approximately .945. Thus Jones sees that, if he is right, it is very likely that he will win the bet.

Smith, on the other hand, believes that p is significantly greater than $\frac{1}{2}$. If he believes that p is as high as .9, we see from Figure 14 that the probability of his getting eight or more correct is .930. Then both men will be satisfied by the bet.

Suppose, however, that Smith thinks the value of p is only about .75. Then the probability that he will get eight or more correct and thus win the bet is .526. There is then only an approximately even chance that the experiment will discover his abilities, and he probably will not be satisfied with this. If Smith really thinks his ability is represented by a p value of about $\frac{3}{4}$, we would have to devise a different method of awarding the dollar. We might, for example, propose that Smith win the bet if he gets seven or more correct. Then, if he has probability $\frac{3}{4}$ of being correct on a single trial, the probability that he will win the bet is approximately .776. If $p = \frac{1}{2}$, the probability that Jones will win the bet is about .828 under this new arrangement. Jones's chances of winning are thus decreased, but Smith may be able to convince him that it is a fairer arrangement than the first procedure.

In the above example, it was possible to make two kinds of errors. The probability of making these errors depended on the way we designed the experiment and the method we used for the required decision. In some cases we are not too worried about the errors and can make a relatively simple experiment. In other cases, errors are very important, and the experiment must be designed with that fact in mind. For example, the possibility of error is certainly important in the case that a vaccine for a given disease is proposed, and the statistician is asked to help in deciding whether or not it should be used. In this case it might be assumed that there is a certain probability p that a person will get the disease if not vaccinated, and a probability r that he will get it if he is vaccinated. If we have some knowledge of the approximate value of p, we are then led to construct an experiment to decide whether r is greater than p, equal to p, or less than p. The first case would be interpreted to mean that the vaccine actually tends to produce the disease, the second that it has no effect, and the third that it prevents the disease; so that we can make three kinds of errors. We could recommend acceptance when

it is actually harmful, we could recommend acceptance when it has no effect, or finally we could reject it when it actually is effective. The first and third might result in the loss of lives, the second in the loss of time and money of those administrating the test. Here it would certainly be important that the probability of the first and third kinds of errors be made small. To see how it is possible to make the probability of both errors small, we return to the case of Smith and Jones.

Suppose that, instead of demanding that Smith make at least eight correct identifications out of ten trials, we insist that he make at least 60 correct identifications out of 100 trials. (The glasses must now be very small.) Then, if $p = \frac{1}{2}$, the probability that Jones wins the bet is about .98; so that we are extremely unlikely to give the dollar to Smith when in fact it should go to Jones. (If $p < \frac{1}{2}$, it is even more likely that Jones will win.) If $p > \frac{1}{2}$, we can also calculate the probability that Smith will win the bet. These probabilities are shown in the graph in Figure 15. The dashed curve gives for com-

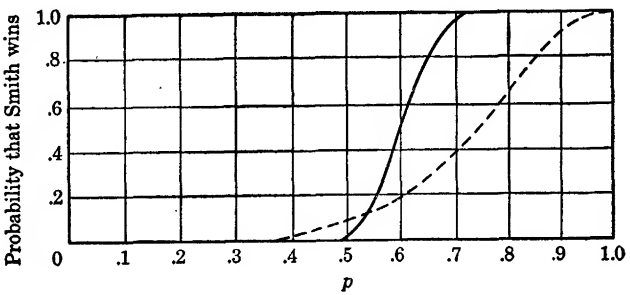


Figure 15

parison the corresponding probabilities for the test requiring eight out of ten correct. Note that with 100 trials, if p is $\frac{3}{4}$, the probability that Smith wins the bet is nearly 1, while in the case of eight out of ten, it was only about $\frac{1}{2}$. Thus in the case of 100 trials, it would be easy to convince both Smith and Jones that whichever one is correct is very likely to win the bet.

Thus we see that the probability of both types of errors can be made small at the expense of having a large number of experiments.

EXERCISES

1. Assume that in the beer and ale experiment Jones agrees to pay Smith if Smith gets at least nine out of ten correct.

(a) What is the probability of Jones paying Smith even though Smith cannot distinguish beer and ale, and guesses? [*Ans.* .011.]

(b) Suppose that Smith can distinguish with probability .9. What is the probability of his not collecting from Jones? [*Ans.* .264.]

2. Suppose that in the beer and ale experiment Jones wishes the probability to be less than .1 that Smith will be paid if, in fact, he guesses. How many of ten trials must he insist that Smith get correct to achieve this?

3. In the analysis of the beer and ale experiment, we assume that the various trials were independent. Discuss several ways that error can enter, because of the nonindependence of the trials, and how this error can be eliminated. (For example, the glasses in which the beer and ale were served might be distinguishable.)

4. Consider the following two procedures for testing Smith's ability to distinguish beer from ale.

(a) Four glasses are given at each trial, three containing beer and one ale, and he is asked to pick out the one containing ale. This procedure is repeated ten times. He must guess correctly seven or more times.

(b) Ten glasses are given him, and he is told that five contain beer and five ale, and he is asked to name the five which he believes contain ale. He must choose all five correctly.

In each case, find the probability that Smith establishes his claim by guessing. Is there any reason to prefer one test over the other?

[*Ans.* (a) .003; (b) .004.]

5. A testing service claims to have a method for predicting the order in which a group of freshmen will finish in their scholastic record at the end of college. The college agrees to try the method on a group of five students, and says that it will adopt the method if, for these five students, the prediction is either exactly correct or can be changed into the correct order by interchanging one pair of *adjacent* men in the predicted order. If the method is equivalent to simply guessing, what is the probability that it will be accepted? [*Ans.* $\frac{1}{24}$.]

6. The standard treatment for a certain disease leads to a cure in $\frac{1}{4}$ of the cases. It is claimed that a new treatment will result in a cure in $\frac{3}{4}$ of the cases. The new treatment is to be tested on ten people having the disease.

If seven or more are cured the new treatment will be adopted. If three or fewer people are cured, the treatment will not be considered further. If the number cured is four, five, or six, the results will be called inconclusive, and a further study will be made. Find the probabilities for each of these three alternatives under the assumption first, that the new treatment has the same effectiveness as the old, and second, under the assumption that the claim made for the treatment is correct.

7. Three students debate the intelligence of blonde dates. One claims that blondes are mostly (say 90 per cent of them) intelligent. A second claims that very few (say ten per cent) blondes are intelligent, while a third one claims that a blonde is just as likely to be intelligent as not. They administer an intelligence test to ten blondes, classifying them as intelligent or not. They agree that the first man wins the bet if eight or more are intelligent, the second if two or fewer, the third in all other cases. For each man, calculate the probability that he wins the bet, if he is right.

[*Ans.* .930, .930, .890.]

8. Ten men take a test with ten problems. Each man on each question has probability $\frac{1}{2}$ of being right, if he does not cheat. The instructor determines the number of students who get each problem correct. If he finds on four or more problems there are fewer than three or more than seven correct, he considers this convincing evidence of communication between the students. Give a justification for the procedure. [*Hint:* The table in Figure 14 must be used twice, once for the probability of fewer than three or more than seven correct answers on a given problem, and the second time to find the probability of this happening on four or more problems.]

*10. THE LAW OF LARGE NUMBERS

In this section we shall study some further properties of the independent trials process with two outcomes. In Section 8 we saw that the probability for x successes in n trials is given by

$$f(n,x;p) = \binom{n}{x} p^x q^{n-x}.$$

In Figure 16 we show these probabilities graphically for $n = 8$ and $p = \frac{3}{4}$. In Figure 17 we have done similarly for the case of $n = 7$ and $p = \frac{3}{4}$.

We see in the first case that the values increase up to a maximum value at $x = 6$ and then decrease. In the second case the values increase up to a maximum value at $x = 5$, have the same value for

$x = 6$, and then decrease, and these two cases are typical of what can happen in general.

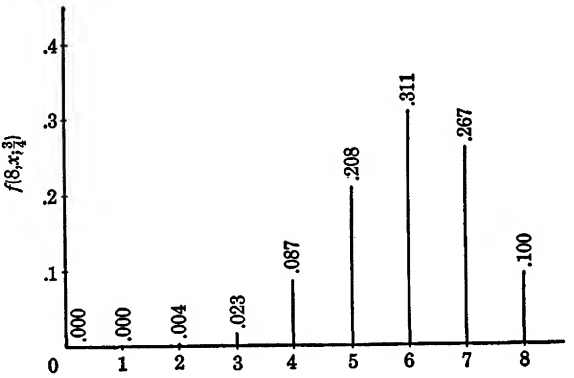


Figure 16

Consider the ratio of the probability of $x + 1$ successes in n trials to the probability of x successes in n trials, which is

$$\frac{\binom{n}{x+1}p^{x+1}q^{n-x-1}}{\binom{n}{x}p^xq^{n-x}} = \frac{n-x}{x+1}\cdot\frac{p}{q}.$$

This ratio will be greater than one as long as $(n - x)p > (x + 1)q$ or as long as $x < np - q$. If $np - q$ is not an integer, the values

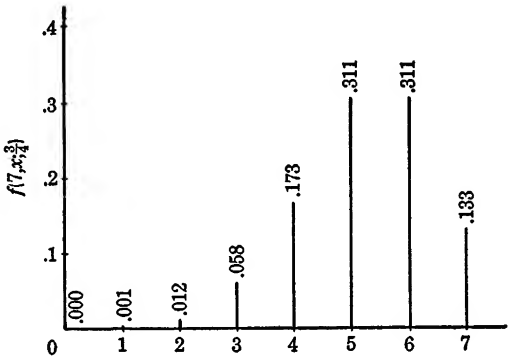


Figure 17

$\binom{n}{x}p^x q^{n-x}$ increase up to a maximum value, which occurs at the first integer greater than $np - q$, and then decrease. In case $np - q$ is an integer, the values $\binom{n}{x}p^x q^{n-x}$ increase up to $x = np - q$, are the same for $x = np - q$ and $x = np - q + 1$, and then decrease.

Thus we see that, in general, values near np will occur with the largest probability. It is not true that one particular value near np is highly likely to occur, but only that it is relatively more likely than a value further from np. For example, in 100 throws of a coin, $np = 100 \cdot \frac{1}{2} = 50$. The probability of exactly 50 heads is approximately .08. The probability of exactly 30 is approximately .00002.

More information is obtained by studying the probability of a given deviation of the proportion of successes x/n from the number p; that is, by studying for ϵ greater than zero, the probability $\Pr[p - \epsilon < x/n < p + \epsilon]$.

For any fixed n, p, and ϵ, the latter probability could be found by adding all the values of $f(n,x;p)$ for values of x for which the inequality $p - \epsilon < x/n < p + \epsilon$ is satisfied. This would, for any particular choice of n, p, and ϵ, be a tedious task. However, it is proved in more advanced books that

$$\Pr[p - \epsilon < \frac{x}{n} < p + \epsilon] \geq 1 - \frac{pq}{n\epsilon^2}.$$

No matter how small ϵ is, if we choose n large enough, we can make $1 - \frac{pq}{n\epsilon^2}$ as near to 1 as we wish. Thus the probability for the proportion of successes deviating from p by less than ϵ can be made arbitrarily near to 1 by choosing n large enough. The fact that this can be done is a special case of a very general theorem in probability theory called the *law of large numbers.*

Let us put in the above inequality $\epsilon = k\sqrt{pq/n}$. Then we have

$$\Pr\left[p - k\sqrt{\frac{pq}{n}} < \frac{x}{n} < p + k\sqrt{\frac{pq}{n}}\right] \geq 1 - \frac{1}{k^2}$$

or

$$\Pr[np - k\sqrt{npq} < x < np + k\sqrt{npq}] \geq 1 - \frac{1}{k^2}$$

which in turn can be written

$$\Pr[-k\sqrt{npq} < x - np < k\sqrt{npq}] \geq 1 - \frac{1}{k^2}.$$

The quantity np is called the *expected value* for the number of successes, and the quantity $\sqrt{npq}$ is called the *standard deviation* for the number of successes. We note that the probability of a deviation of more than k standard deviations from the expected value is less than or equal to $1/k^2$. Thus for large k this probability will be small.

It is possible to show by the more advanced theory that

$$\Pr[-k\sqrt{npq} < x - np < k\sqrt{npq}] \cong z_k$$

where z_k is a number which can be found for any k and does not depend on n or on p. The symbol $\cong$ means that the indicated probability is only approximately given by z_k. The approximation improves with increasing n, but the error may be significant even for reasonably large n, and hence in practice this approximation must be used with care.

We note that the approximation given above can also be interpreted as stating that the probability that $x - np$ is *either* greater than $k\sqrt{npq}$ or less than $-k\sqrt{npq}$ is approximately $1 - z_k$. In many applications one is interested only in the probability that $x - np$ is greater than $k\sqrt{npq}$. It follows from the more advanced theory that this is approximately $(1 - z_k)/2$. Hence also the probability that $x - np$ is less than $-k\sqrt{npq}$ is approximately $(1 - z_k)/2$.

It is convenient to think of the standard deviation as a unit of measurement. In this case z_k gives the approximate probability for a deviation of less than k units, or k standard deviations. The value of z_k for $k = 1$, 2, and 3 are $z_1 = .683 \ldots$, $z_2 = .956 \ldots$, $z_3 = .997 \ldots$. Thus we see that it is very unlikely in a large number of trials to have a deviation from the expected value of more than three standard deviations. On the other hand $z_{.1} = .080 \ldots$ which shows that it is quite unlikely that there will be a deviation of less than one-tenth of a standard deviation from the expected value.

Example 1. In throwing an ordinary coin 10,000 times, the expected number of heads is 5000, and the standard deviation for the number of heads is $\sqrt{10{,}000(\frac{1}{2})(\frac{1}{2})} = 50$. Thus the probability that the number of heads which turn up deviates from 5000 by less than one standard deviation, or 50, is approximately .683. The probability of a deviation of less than two standard deviations, or 100, is approximately .954. The probability of a deviation of less than three stand-

ard deviations, or 150, is approximately .997. On the other hand, the probability of a deviation of less than .1 standard deviation, or a deviation of less than five, is approximately .080. The statement that the number of heads deviates from 5000 by less than 150 is equivalent to the statement that the proportion of heads deviates from .5 by less than $150/10{,}000 = .015$.

Example 2. Assume that in a certain large city, 900 people are chosen at random and asked if they favor a certain proposal. Of the 900 asked, 550 say they favor the proposal and 350 are opposed. If, in fact, the people in the city are equally divided on the issue, would it be unlikely that such a large majority would be obtained in a sample of 900 of the citizens? If the people were equally divided, we would assume that the 900 people asked would form an independent trials process with probability $\frac{1}{2}$ for a "yes" answer and $\frac{1}{2}$ for a "no" answer. Then the standard deviation for the number of "yes" answers in 900 trials is $\sqrt{900(\frac{1}{2})(\frac{1}{2})} = 15$. Then it would be very unlikely that we would obtain a deviation of more than 45 from the expected number of 450. The fact that the deviation in the sample from the expected number was 100, then, is evidence that the hypothesis that the voters were equally divided is incorrect. The assumption that the true proportion is any value less than $\frac{1}{2}$ would also lead to the fact that a number as large as 550 favoring in a sample of 900 is very unlikely. Thus we are led to suspect that the true proportion is greater than $\frac{1}{2}$. On the other hand, if the number who favored the proposal in the sample of 900 were 465, we would have only a deviation of one standard deviation, under the assumption of an equal division of opinion. Since such a deviation is not unlikely, we could not rule out this possibility on the evidence of the sample.

EXERCISES

1. If an ordinary die is thrown 20 times, what is the expected number of times that a 6 will turn up? What is the standard deviation for the number of 6's that turn up? [*Ans.* $\frac{10}{3}$; $\frac{5}{3}$.]

2. Suppose that an ordinary die is thrown 450 times. What is the expected number of throws that result in either a 3 or a 4? What is the standard deviation for the number of such throws?

3. In 16 tosses of an ordinary coin, what is the expected number of heads that turn up? What is the standard deviation for the number of heads that occur? [*Ans.* 8; 2.]

4. In 16 tosses of a coin, find the exact probability that the number of heads that turn up differs from the expected number by (a) less than one standard deviation, and (b) by not more than one standard deviation. Do the same for the case of two standard deviations, and for the case of three standard deviations. Show that the approximations given for large n lie between the values obtained, but are not very accurate for so small an n. [*Ans.* .546, .790; .923, .979; .996, .999.]

5. Consider n independent trials with probability p for success. Let r and s be numbers such that $p < r < s$. What does the law of large numbers say about

$$\Pr\left[r < \frac{x}{n} < s\right]$$

as we increase n indefinitely? Answer the same question in the case that $r < p < s$.

6. A drug is known to be effective in 20 per cent of the cases where it is used. A new agent is introduced, and in the next 900 times the drug is used it is effective 250 times. What can be said about the effectiveness of the drug?

7. In a large number of independent trials with probability p for success, what is the approximate probability that the number of successes will deviate from the expected number by more than one standard deviation but less than two standard deviations? [*Ans.* .271.]

8. What is the approximate probability that, in 10,000 throws of an ordinary coin, the number of heads which turn up lies between 4850 and 5150? What is the probability that the number of heads lies in the same interval, given that in the first 1900 throws there were 1600 heads?

9. Suppose that it is desired that the probability be approximately .95 that the fraction of 6's that turn up when a die is thrown n times does not deviate by more than .01 from the value $\frac{1}{6}$. How large should n be? [*Ans.* Approximately 5555.]

10. Two railroads are competing for the passenger traffic of 1000 passengers by operating similar trains at the same hour. If a given passenger is equally likely to choose one train as the other, how many seats should the railroad provide if it wants to be sure that its seating capacity is sufficient in 99 out of 100 cases? [*Ans.* 547.]

11. Assume that 10 per cent of the people in a certain city have cancer. If 900 people are selected at random from the city, what is the expected

number which will have cancer? What is the standard deviation? What is the approximate probability that more than 108 of the 900 chosen have cancer? [*Ans.* 90; 9; .023.]

12. Suppose that in Exercise 11, the 900 people are chosen at random from those people in the city who smoke. Under the hypothesis that smoking has no effect on the incidence of cancer, what is the expected number in the 900 chosen that have cancer? Suppose that more than 120 of the 900 chosen have cancer, what might be said concerning the hypothesis that smoking has no effect on the incidence of cancer?

13. In Example 2, we made the assumption in our calculations that, if the true proportion of voters in favor of the proposal were p, then the 900 people chosen at random represented an independent trials process with probability p for a "yes" answer, and $1 - p$ for a "no" answer. Give a method for choosing the 900 people which would make this a reasonable assumption. Criticize the following methods:

(a) Choose the first 900 people in the list of registered Republicans.
(b) Choose 900 names at random from the telephone book.
(c) Choose 900 houses at random and ask one person from each house, the houses being visited in the mid-morning.

14. For n throws of an ordinary coin, let t_n be such that

$$\Pr\left[-t_n < \frac{x}{n} - \frac{1}{2} < t_n\right] = .997$$

where x is the number of heads that turn up. Find t_n for $n = 10^4$, $n = 10^6$, and $n = 10^{20}$. [*Ans.* .015; .0015; .000,000,000,15.]

15. Assume that a calculating machine carries out a million operations to solve a certain problem. In each operation the machine gives the answer 10^{-5} too small, with probability $\frac{1}{2}$, and 10^{-5} too large, with probability $\frac{1}{2}$. Assume that the errors are independent of one another. What is a reasonable accuracy to attach to the answer? What if the machine carries out 10^9 operations? [*Ans.* $\pm.01$; ±1.]

*11. INDEPENDENT TRIALS WITH MORE THAN TWO OUTCOMES

By extending the results of Section 8, we shall study the case of independent trials in which we allow more than two outcomes. We assume that we have an independent trials process where the possible outcomes are $a_1, a_2, \ldots, a_k$, occurring with probabilities $p_1, p_2, \ldots, p_k$, respectively. We denote by $f(r_1,r_2, \ldots, r_k;p_1,p_2, \ldots, p_k)$ the probability that, in $n = r_1 + r_2 + \ldots + r_k$ such trials, there will be

r_1 occurrences of a_1, r_2 of a_2, etc. In the case of two outcomes this notation would be $f(r_1,r_2;p_1,p_2)$. In Section 8 we wrote this as $f(n,r_1;p_1)$ since r_2 and p_2 are determined from n, r_1, and p_1. We shall indicate how this probability is found in general, but carry out the details only for a special case. We choose $k = 3$, and $n = 5$ for purposes of illustration. We shall find $f(1,2,2;p_1,p_2,p_3)$.

We show in Figure 18 enough of the tree for this process to indicate the branch probabilities for a path (heavy lined) corresponding to the outcomes a_2, a_3, a_1, a_2, a_3. The tree measure assigns weight $p_2 \cdot p_3 \cdot p_1 \cdot p_2 \cdot p_3 = p_1 \cdot p_2^2 \cdot p_3^2$ to this path.

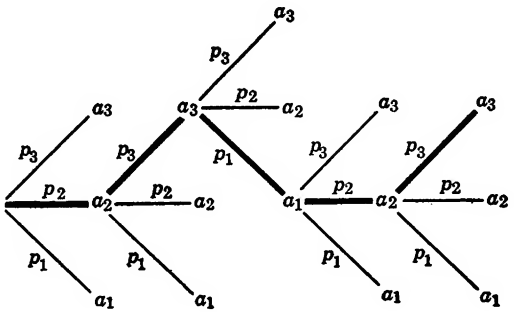


Figure 18

There are, of course, other paths through the tree corresponding to one occurrence of a_1, two of a_2 and two of a_3. However, they would all be assigned the same weight, $p_1 \cdot p_2^2 \cdot p_3^2$, by the tree measure. Hence to find $f(1,2,2;p_1,p_2,p_3)$, we must multiply this weight by the number of paths having the specified number of occurrences of each outcome.

We note that the path a_2, a_3, a_1, a_2, a_3 can be specified by the three-cell partition $[\{3\}, \{1,4\}, \{2,5\}]$ of the numbers from 1 to 5. Here the first cell shows the experiment which resulted in a_1, the second cell shows the two that resulted in a_2, and the third shows the two that resulted in a_3. Conversely, any such partition of the numbers from 1 to 5 with one element in the first cell, two in the second, and two in the third corresponds to a unique path of the desired kind. Hence the number of paths is the number of such partitions. But this is

$$\binom{5}{1,2,2} = \frac{5!}{1!2!2!}$$

(see Chapter 3, Section 5), so that the probability of one occurrence of a_1, two of a_2, and two of a_3 is

$$\binom{5}{1,2,2} \cdot p_1 \cdot p_2^2 \cdot p_3^2.$$

The above argument carried out in general leads, for the case of independent trials with outcomes $a_1, a_2, \ldots, a_k$ occurring with probabilities $p_1, p_2, \ldots, p_k$, to the following.

The probability for r_1 occurrences of a_1, r_2 occurrences of a_2, etc., is given by

$$f(r_1,r_2, \ldots, r_k;p_1,p_2, \ldots, p_k) = \binom{n}{r_1,r_2, \ldots, r_k} p_1^{r_1} \cdot p_2^{r_2} \cdots p_k^{r_k}.$$

Example 1. A die is thrown 12 times. What is the probability that each number will come up twice? Here there are six outcomes, 1, 2, 3, 4, 5, 6 corresponding to the six sides of the die. We assign each outcome probability $\frac{1}{6}$. We are then asked for

$$f(2,2,2,2,2,2;\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6},\tfrac{1}{6})$$

which is

$$\binom{12}{2,2,2,2,2,2}\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^2 = .0034 \ldots .$$

Example 2. Suppose that we have a repeated-trials process with four outcomes a_1, a_2, a_3, a_4 occurring with probability p_1, p_2, p_3, p_4, respectively. It might be that we are interested only in the probability that r_1 occurrences of a_1 and r_2 occurrences of a_2 will take place with no specification about the number of each of the other possible outcomes. To answer this question we simply consider a new experiment where the outcomes are a_1, a_2, $\bar{a}_3$. Here $\bar{a}_3$ corresponds to an occurrence of either a_3 or a_4 in our original experiment. The corresponding probabilities would be p_1, p_2, and $\bar{p}_3$ with $\bar{p}_3 = p_3 + p_4$. Let $\bar{r}_3 = n - (r_1 + r_2)$. Then our question is answered by finding the probability in our new experiment for r_1 occurrences of a_1, r_2 of a_2, and $\bar{r}_3$ of $\bar{a}_3$, which is

$$\binom{n}{r_1,r_2,\bar{r}_3} p_1^{r_1} \cdot p_2^{r_2} \cdot p_3^{\bar{r}_3}.$$

The same procedure can be carried out for experiments with any

number of outcomes where we specify the number of occurrences of such particular outcomes. For example, if a die is thrown ten times the probability that a one will occur exactly twice and a three exactly three times is given by

$$\binom{10}{2,3,5}\left(\frac{1}{6}\right)^2\left(\frac{1}{6}\right)^3\left(\frac{4}{6}\right)^5 = .043 \ldots .$$

EXERCISES

1. Suppose that in a city 60 per cent of the population are Democrats, 30 per cent are Republicans, and 10 per cent are Independents. What is the probability that if three people are chosen at random there will be one Republican, one Democrat, and one Independent voter? [*Ans.* .108.]

2. Three horses A, B, and C compete in four races. Assuming that each horse has an equal chance in each race, what is the probability that A wins two races and B and C win one each? What is the probability that the same horse wins all four races? [*Ans.* $\frac{4}{27}$; $\frac{1}{27}$.]

3. Assume that in a certain large college 40 per cent of the students are freshmen, 30 per cent are sophomores, 20 per cent are juniors, and 10 per cent are seniors. A committee of eight is chosen at random from the student body. What is the probability that there are equal numbers from each class on the committee?

4. Let us assume that when a batter comes to bat, he has probability .6 of being put out, .1 of getting a walk, .2 of getting a single, .1 of getting an extra base hit. If he comes to bat five times in a game, what is the probability that

(a) He gets two walks and three singles? [*Ans.* .0008.]

(b) A walk, a single, an extra base hit (and is out twice)? [*Ans.* .043.]

(c) Has a perfect day (i.e., never out). [*Ans.* .010.]

5. Assume that a single torpedo has a probability $\frac{1}{2}$ of sinking a ship, probability $\frac{1}{4}$ of damaging it, and probability $\frac{1}{4}$ of missing. Assume further that two damaging shots sink the ship. What is the probability that four torpedos will succeed in sinking the ship? [*Ans.* $\frac{25}{256}$.]

6. Jones, Smith, and Green live in the same house. The mailman has observed that Jones and Smith receive the same amount of mail on the average, but that Green receives twice as much as Jones (and hence also twice as much as Smith). If he has four letters for this house, what is the probability that each man receives at least one letter?

7. If three dice are thrown, find the probability that there is one six and two fives, given that all the outcomes are greater than three. [*Ans.* $\frac{1}{9}$.]

8. A man plays a tournament consisting of three games. In each game he has probability $\frac{1}{2}$ for a win, $\frac{1}{4}$ for a loss, and $\frac{1}{4}$ for a draw, independently of the outcomes of other games. To win the tournament he must win more games than he loses. What is the probability that he wins the tournament?

9. Assume that in a certain course the probability that a student chosen at random will get an A is .1, that he will get a B is .2, that he will get a C is .4, that he will get a D is .2, and that he will get an E is .1. What distribution of grades is most likely in the case of four students?

[*Ans.* One B, two C's, one D.]

10. Let us assume that in a World Series game a batter has probability $\frac{1}{4}$ of getting no hits, $\frac{1}{2}$ for getting one hit, $\frac{1}{4}$ for getting two hits, assuming that the probability of getting more than two hits is negligible. In a four-game World Series, find the probability that the batter gets:

(a) Exactly two hits.
(b) Exactly three hits.
(c) Exactly four hits.
(d) Exactly five hits.
(e) Fewer than two hits or more than five.

[*Ans.* $\frac{7}{64}$, $\frac{7}{32}$, $\frac{35}{128}$, $\frac{7}{32}$, $\frac{23}{128}$.]

12. EXPECTED VALUE

In this section we shall discuss the concept of expected value. Although it originated in the study of gambling games, it enters into almost any detailed probabilistic discussion.

DEFINITION. If in an experiment the possible outcomes are numbers, $a_1, a_2, \ldots, a_k$, occurring with probability $p_1, p_2, \ldots, p_k$, then the *expected value* is defined to be

$$E = a_1p_1 + a_2p_2 + \ldots + a_kp_k.$$

The term "expected value" is not to be interpreted as the value that will necessarily occur on a single experiment. For example, if a person bets \$1 that a head will turn up when a coin is thrown, he expects to win \$1 or to lose \$1. His expected value is $(1)(\frac{1}{2}) + (-1)(\frac{1}{2}) = 0$, which is not one of the possible outcomes. The term, expected value, had its origin in the following consideration. If we repeat an experiment with expected value E a large number of times, and if we expect

a_1 a fraction p_1 of the time, a_2 a fraction p_2 of the time, etc., then the average that we expect per experiment is E. In particular, in a gambling game E is interpreted as the average winning expected in a large number of plays. Here the expected value is often taken as the value of the game to the player. If the game has a positive expected value, the game is said to be favorable, if the game has expected value zero it is said to be fair, and if it has negative expected value it is described as unfavorable. These terms are not to be taken too literally, since many people are quite happy to play games that, in terms of expected value, are unfavorable.

Example 1. For the first example of the application of expected value we consider the game of roulette as played at Monte Carlo. There are several types of bets which the gambler can make, and we consider two of these.

The wheel has the number 0 and the numbers from 1 to 36 marked on equally spaced slots. The wheel is spun and a ball comes to rest in one of these slots. If the player puts a stake, say of \$1, on a given number, and the ball comes to rest in this slot, then he receives from the croupier 36 times his stake, or \$36. Thus for a payment of \$1 his expected winning is $\frac{36}{37} = .973$. This can be interpreted to mean that in the long run he can expect to lose about 2.7 per cent of his stakes.

A second way to play is the following. A player may bet on "red" or "black." The numbers from 1 to 36 are evenly divided between the two colors. If a player bets on "red," and a red number turns up, he receives twice his stake. If a black number turns up, he loses his stake. If 0 turns up, then the wheel is spun until it stops on a number different from 0. If this is black, the player loses; but if it is red, he receives only his original stake, not twice it. For this type of play, the gambler pays \$1 for an expected winning of

$$2(\tfrac{18}{37}) + 1(\tfrac{1}{74}) = \tfrac{73}{74} = .9865.$$

In this case the player can expect to lose about 1.35 per cent of his stakes in the long run. Thus the expected loss in this case is only half as great as in the previous case.

Example 2. A player rolls a die and receives a number of dollars corresponding to the number of dots on the face which turns up. What should the player pay for playing, to make this a fair game?

To answer this question, we note that the player wins 1, 2, 3, 4, 5 or 6 dollars, each with probability $\frac{1}{6}$. Hence, his expected winning is

$$1(\tfrac{1}{6}) + 2(\tfrac{1}{6}) + 3(\tfrac{1}{6}) + 4(\tfrac{1}{6}) + 5(\tfrac{1}{6}) + 6(\tfrac{1}{6}) = 3\tfrac{1}{2}.$$

Thus if he pays \$3.50, his expected winnings will be zero.

Example 3. What is the expected number of successes in the case of four independent trials with probability $\frac{1}{3}$ for success? We know that the probability of x successes is $\binom{4}{x}\left(\frac{1}{3}\right)^x\left(\frac{2}{3}\right)^{4-x}$. Thus

$$E = 0\cdot\binom{4}{0}(\tfrac{1}{3})^0(\tfrac{2}{3})^4 + 1\cdot\binom{4}{1}(\tfrac{1}{3})^1(\tfrac{2}{3})^3 + 2\cdot\binom{4}{2}(\tfrac{1}{3})^2(\tfrac{2}{3})^2$$
$$+ 3\cdot\binom{4}{3}(\tfrac{1}{3})^3(\tfrac{2}{3})^1 + 4\cdot\binom{4}{4}(\tfrac{1}{3})^4(\tfrac{2}{3})^0$$
$$= 0 + \tfrac{32}{81} + \tfrac{48}{81} + \tfrac{24}{81} + \tfrac{4}{81} = \tfrac{108}{81} = \tfrac{4}{3}.$$

In general, it can be shown that in n trials with probability p for success, the expected number of successes is np.

Example 4. In the game of craps a pair of dice is rolled by one of the players. If the sum of the spots shown is 7 or 11, he wins. If it is 2, 3, or 12, he loses. If it is another sum, he must continue rolling the dice until he either repeats the same sum or rolls a 7. In the former case he wins, in the latter he loses. Let us suppose that he wins or loses \$1. Then the two possible outcomes are $+1$ and -1. We will compute the expected value of the game. First we must find the probability that he will win.

We represent the possibilities by a two-stage tree shown in Figure 19. While it is theoretically possible for the game to go on indefinitely, we do not consider this possibility. This means that our analysis applies only to games which actually stop at some time.

The branch probabilities at the first stage are determined by thinking of the 36 possibilities for the throw of the two dice as being equally likely and taking in each case the fraction of the possibilities which correspond to the branch as the branch probability. The probabilities for the branches at the second level are obtained as follows. If, for example, the first outcome was a 4, then when the game ends, a 4 or 7 must have occurred. The possible outcomes for the dice were {(3,1), (1,3), (2,2), (4,3), (3,4), (2,5), (5,2), (1,6), (6,1)}. Again we

consider these possibilities to be equally likely and assign to the branch considered the fraction of the outcomes which correspond to this branch. Thus to the 4 branch we assign a probability $\frac{3}{9} = \frac{1}{3}$. The other branch probabilities are determined in a similar way.

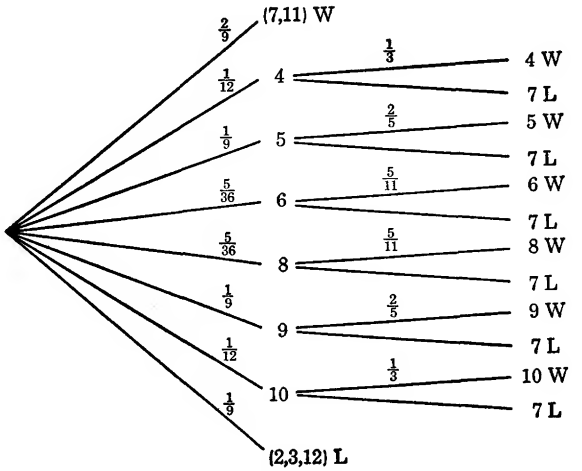


Figure 19

Having the tree measure assigned, to find the probability of a win we must simply add the weights of all paths leading to a win. If this is done, we obtain $\frac{244}{495}$. Thus the player's expected value is $1 \cdot (\frac{244}{495}) + (-1) \cdot (\frac{251}{495}) = -\frac{7}{495} = -.0141$. Hence he can expect to lose 1.41 per cent of his stakes in the long run. It is interesting to note that this is just slightly less favorable than his losses in betting on "red" in roulette.

EXERCISES

1. Suppose that A tosses 2 coins and receives \$2 if two heads appear, \$1 if one head appears, and nothing if no heads appear. What is the expected value of the game to him? [*Ans.* \$1.]

2. Smith and Jones are matching coins. If the coins match, Smith gets \$1, and if they do not, Jones gets \$1.

(a) If the game consists of matching twice, what is the expected value of the game for Smith?

(b) Suppose that if Smith wins the first round he quits, and if he loses the first he plays the second. Jones is not allowed to quit. What is the expected value of the game for Smith?

3. If five coins are thrown, what is the expected number of heads that will turn up? [*Ans.* $\frac{5}{2}$.]

4. A coin is thrown until the first time a head comes up or until three tails in a row occur. Find the expected number of times the coin is thrown.

5. A man wishes to purchase a five cent newspaper. He has in his pocket one dime and five pennies. The newsman offers to let him have the paper in exchange for one coin drawn at random from the customer's pocket.

(a) Is this a fair proposition and, if not, to whom is it favorable? [*Ans.* Favorable to man.]

(b) Answer the same questions as in (a) assuming that the newsman demands two coins drawn at random from the customer's pocket. [*Ans.* Fair proposition.]

6. A bets 50 cents against B's x cents that, if two cards are dealt from a shuffled pack of ordinary playing cards, both cards will be of the same color. What value of x will make this bet fair?

7. Prove that if the expected value of a given experiment is E, and if a constant c is added to each of the outcomes, the expected value of the new experiment is $E + c$.

8. Prove that, if the expected value of a given experiment is E, and if each of the possible outcomes is multiplied by a constant k, the expected value of the new experiment is $k \cdot E$.

9. Referring to Example 2, Section 7, find the expected number of blocks the man will walk before reaching home. [*Ans.* $\frac{5}{2}$.]

10. An urn contains two black and three white balls. Balls are successively drawn from the urn without replacement until a black ball is obtained. Find the expected number of draws required.

11. Using the result of Exercises 15, 16 of Section 7, find the expected number of games in the World Series (a) under the assumption that each team has probability $\frac{1}{2}$ of winning each game and (b) under the assumption that the stronger team has probability $\frac{2}{3}$ of winning each game. [*Ans.* 5.81; 5.50.]

12. Suppose that we modify the game of craps as follows: On a 7 or 11 the player wins \$2, on a 2, 3, or 12 he loses \$3; otherwise the game is as usual. Find the expected value of the new game, and compare it with the old value.

13. Suppose that in roulette at Monte Carlo we place 50 cents on "red" and 50 cents on "black." What is the expected value on the game? Is this better or worse than placing \$1 on "red"?

14. Betting on "red" in roulette can be described roughly as follows. We win with probability .49, get our money back with probability .01, and lose with probability .50. Draw the tree for three plays of the game, and compute (to three decimals) the probability of each path. What is the probability that we are ahead at the end of three bets? [*Ans.* .485.]

15. Assume that the odds are $r:s$ that a certain statement will be true. If a man receives s dollars if the statement turns out to be true, and gives r dollars if not, what is his expected winning?

16. Referring to Exercise 9 of Section 3, find the expected number of languages that a student chosen at random reads.

17. Referring to Exercise 5 of Section 4, find the expected number of men who get their own hats. [*Ans.* 1.]

13. MARKOV CHAINS

In this section we shall study a more general kind of process than the ones considered in the last three sections.

We assume that we have a sequence of experiments with the following properties. The outcome of each experiment is one of a finite number of possible outcomes $a_1, a_2, \ldots, a_r$. It is assumed that the probability of outcome a_j on any given experiment is not necessarily independent of the outcomes of previous experiments but depends at most upon the outcome of the immediately preceding experiment. We assume that there are given numbers p_{ij} which represent the probability of outcome a_j on any given experiment, given that outcome a_i occurred on the preceding experiment. The outcomes $a_1, a_2, \ldots, a_r$ are called *states*, and the numbers p_{ij} are called *transition probabilities*. If we assume that the process begins in some particular state, then we have enough information to determine the tree measure for the process and can calculate probabilities of statements relating to the over-all sequence of experiments. A process of the above kind is called a *Markov chain process.*

The transition probabilities can be exhibited in two different ways.

The first way is that of a square array. For a Markov chain with states a_1, a_2, and a_3, this array is written as

$$P = \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix}.$$

Such an array is a special case of a *matrix*. Matrices are of fundamental importance to the study of Markov chains as well as being important in the study of other branches of mathematics. They will be studied in detail in the next chapter.

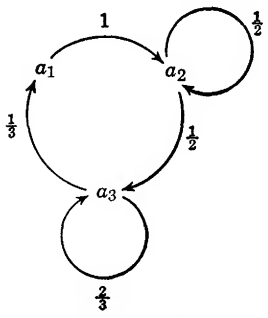


Figure 20

A second way to show the transition probabilities is by a *transition diagram*. Such a diagram is illustrated for a special case in Figure 20. The arrows from each state indicate the possible states to which a process can move from the given state.

The matrix of transition probabilities which corresponds to this diagram is the matrix

$$P = \begin{array}{c} \\ a_1 \\ a_2 \\ a_3 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix} \end{array}.$$

An entry of 0 indicates that the transition is impossible.

Notice that in the matrix P the sum of the elements of each row is 1. This must be true in any matrix of transition probabilities, since the elements of the ith row represent the probabilities for all possibilities when the process is in state a_i.

The kind of problem in which we are most interested in the study of Markov chains is the following. Suppose that the process starts in state i. What is the probability that after n steps it will be in state j? We denote this probability by $p_{ij}^{(n)}$. Notice that we do *not* mean by this the nth power of the number p_{ij}. We are actually interested in this probability for all possible starting positions i and all possible terminal positions j. We can represent these numbers conveniently again by a matrix. For example for n steps in a three-state Markov chain we write these probabilities as the matrix

$$P^{(n)} = \begin{pmatrix} p_{11}^{(n)} & p_{12}^{(n)} & p_{13}^{(n)} \\ p_{21}^{(n)} & p_{22}^{(n)} & p_{23}^{(n)} \\ p_{31}^{(n)} & p_{32}^{(n)} & p_{33}^{(n)} \end{pmatrix}.$$

Example 1. Let us find for a Markov chain with transition probabilities indicated in Figure 20 the probability of being at the various possible states after three steps, assuming that the process starts at state a_1. We find these probabilities by constructing a tree and a tree measure as in Figure 21.

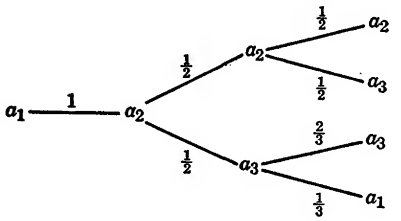


Figure 21

The probability $p_{13}^{(3)}$, for example, is the sum of the weights assigned by the tree measure to all paths through our tree which end at state a_3. That is, $1 \cdot \frac{1}{2} \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} \cdot \frac{2}{3} = \frac{7}{12}$. Similarly $p_{12}^{(3)} = 1 \cdot \frac{1}{2} \cdot \frac{1}{2} = \frac{1}{4}$ and $p_{11}^{(3)} = 1 \cdot \frac{1}{2} \cdot \frac{1}{3} = \frac{1}{6}$. By constructing a similar tree measure, assuming that we start at state a_2, we could find $p_{21}^{(3)}$, $p_{22}^{(3)}$, and $p_{23}^{(3)}$. The same is true for $p_{31}^{(3)}$, $p_{32}^{(3)}$, and $p_{33}^{(3)}$. If this is carried out (see Exercise 7) we can write the results in matrix form as follows:

$$P^{(3)} = \begin{array}{c} \\ a_1 \\ a_2 \\ a_3 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{pmatrix} \frac{1}{6} & \frac{1}{4} & \frac{7}{12} \\ \frac{7}{36} & \frac{7}{24} & \frac{37}{72} \\ \frac{4}{27} & \frac{7}{18} & \frac{25}{54} \end{pmatrix} \end{array}.$$

Again the rows add up to 1, corresponding to the fact that if we start at a given state we must reach some state after three steps. Notice now that all the elements of this matrix are positive, showing that it is possible to reach any state from any state in three steps. In the next chapter we will develop a simple method of computing $P^{(n)}$.

Example 2. Suppose that we are interested in studying the way in which a given state votes in a series of national elections. We wish to make long-term predictions and so will not consider conditions peculiar to a particular election year. We shall base our predictions only on past history of the outcomes of the elections, Republican or Democratic. It is clear that a knowledge of these past results would

influence our predictions for the future. As a first approximation, we assume that the knowledge of the past beyond the last election would not cause us to change the probabilities for the outcomes on the next election. With this assumption we obtain a Markov chain with two states R and D and matrix of transition probabilities

$$\begin{array}{c} \\ R \\ D \end{array} \begin{array}{c} \begin{array}{cc} R & D \end{array} \\ \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \end{array}.$$

The numbers a and b could be estimated from past results as follows. We could take for a the fraction of the previous years in which the outcome has changed from Republican in one year to Democratic in the next year, and for b the fraction of reverse changes.

We can obtain a better approximation by taking into account the previous two elections. In this case our states are RR, RD, DR, and DD, indicating the outcome of two successive elections. Being in state RR means that the last two elections were Republican victories. If the next election is a Democratic victory, we will be in state RD. If the election outcomes for a series of years is $DDDRDRR$, then our process has moved from state DD to DD to DR to RD to DR, and finally to RR. Notice that the first letter of the state to which we move must agree with the second letter of the state from which we came, since these refer to the same election year. Our matrix of transition probabilities will then have the form,

$$\begin{array}{c} \\ RR \\ DR \\ RD \\ DD \end{array} \begin{array}{c} \begin{array}{cccc} RR & DR & RD & DD \end{array} \\ \begin{pmatrix} 1-a & 0 & a & 0 \\ b & 0 & 1-b & 0 \\ 0 & 1-c & 0 & c \\ 0 & d & 0 & 1-d \end{pmatrix} \end{array}.$$

Again the numbers a, b, c, and d would have to be estimated. The study of this example is continued in Chapter V, Section 8.

Example 3. The following example of a Markov chain has been used in physics as a simple model for diffusion of gases. We shall see later that a similar model applies to an idealized problem in changing populations.

We imagine n black balls and n white balls which are put into two urns so that there are n balls in each urn. A single experiment consists in choosing a ball from each urn at random and putting the ball

obtained from the first urn into the second urn, and the ball obtained from the second urn into the first. We take as state the number of black balls in the first urn. If at any time we know this number, then we know the exact composition of each urn. That is, if there are j black balls in urn 1, there must be $n - j$ black balls in urn 2, $n - j$ white balls in urn 1, and j white balls in urn 2. If the process is in state j, then after the next exchange it will be in state $j - 1$, if a black ball is chosen from urn 1 and a white ball from urn 2. It will be in state j if a ball of the same color is drawn from each urn. It will be in state $j + 1$ if a white ball is drawn from urn 1 and a black ball from urn 2. The transition probabilities are then given by (see Exercise 12):

$$
\begin{aligned}
p_{jj-1} &= \left(\frac{j}{n}\right)^2 && j > 0 \\
p_{jj} &= \frac{2j(n-j)}{n^2} \\
p_{jj+1} &= \left(\frac{n-j}{n}\right)^2 && j < n \\
p_{jk} &= 0 && \text{otherwise.}
\end{aligned}
$$

A physicist would be interested, for example, in predicting the composition of the urns after a certain number of exchanges have taken place. Certainly any predictions about the early stages of the process would depend upon the initial composition of the urns. For example, if we started with all black balls in urn 1, we would expect that for some time there would be more black balls in urn 1 than in urn 2. On the other hand, it might be expected that the effect of this initial distribution would wear off after a large number of exchanges. We shall see later, in Chapter V, Section 8, that this is indeed the case.

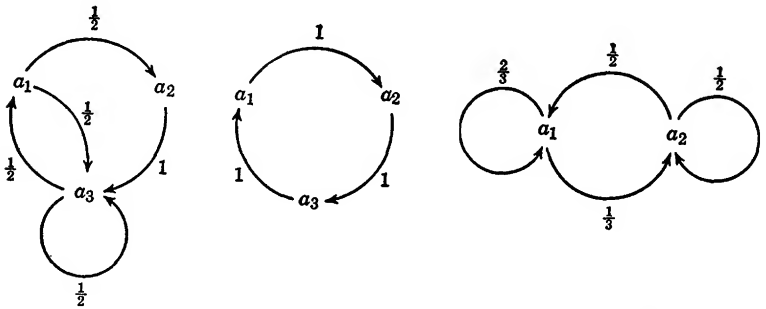
EXERCISES

1. Draw a state diagram for the Markov chain with transition probabilities given by the following matrices.

$$
\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{pmatrix}, \qquad \begin{pmatrix} \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix},
$$

$$
\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \\ 0 & 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix}.
$$

2. Give the matrix of transition probabilities corresponding to the following transition diagrams.



3. Find the matrix $P^{(2)}$ for the Markov chain determined by the matrix of transition probabilities

$$P = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}. \qquad \left[Ans. \begin{pmatrix} \frac{5}{12}, & \frac{7}{12} \\ \frac{7}{18}, & \frac{11}{18} \end{pmatrix}. \right]$$

4. What is the matrix of transition probabilities for the Markov chain in Example 3, for the case of 2 white balls and 2 black balls?

5. Find the matrices $P^{(2)}$, $P^{(3)}$, $P^{(4)}$ for the Markov chain determined by the transition probabilities

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Find the same for the Markov chain determined by the matrix

$$\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

6. Suppose that a Markov chain has two states, a_1 and a_2 and transition probabilities given by the matrix

$$\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

By means of a separate chance device we choose a state in which to start the process. This device chooses a_1 with probability $\frac{1}{2}$ and a_2 with probability $\frac{1}{2}$. Find the probability that the process is in state a_1 after the first step. Answer the same question in the case that the device chooses a_1 with probability $\frac{1}{3}$ and a_2 with probability $\frac{2}{3}$. [*Ans.* $\frac{5}{12}$; $\frac{4}{9}$.]

7. Referring to the Markov chain with transition probabilities indicated in Figure 20, construct the tree measures and determine the values of

$$p_{21}^{(3)}, p_{22}^{(3)}, p_{23}^{(3)}, \quad \text{and} \quad p_{31}^{(3)}, p_{32}^{(3)}, p_{33}^{(3)}$$

8. A certain calculating machine uses only the digits 0 and 1. It is supposed to transmit one of these digits through several stages. However, at every stage there is a probability p that the digit which enters this stage will be changed when it leaves. We form a Markov chain to represent the process of transmission by taking as states the digits 0 and 1. What is the matrix of transition probabilities?

9. For the Markov chain in Exercise 8, draw a tree and assign a tree measure, assuming that the process begins in state 0 and moves through three stages of transmission. What is the probability that the machine after three stages produces the digit 0, i.e., the correct digit? What is the probability that the machine never changed the digit from 0?

10. Assume that a man's profession can be classified as professional, skilled laborer, or unskilled laborer. Assume that of the sons of professional men 80 per cent are professional, 10 per cent are skilled laborers, and 10 per cent are unskilled laborers. In the case of sons of skilled laborers, 60 per cent are skilled laborers, 20 per cent are professional, and 20 per cent are unskilled laborers. Finally, in the case of unskilled laborers, 50 per cent of the sons are unskilled laborers, and 25 per cent each are in the other two categories. Assume that every man has a son, and form a Markov chain by following a given family through several generations. Set up the matrix of transition probabilities. Find the probability that the grandson of an unskilled laborer is a professional man. [*Ans.* .375.]

11. In Exercise 10 we assumed that every man has a son. Assume instead that the probability a man has a son is .8. Form a Markov chain with four states. The first three states are as in Exercise 10, and the fourth state is such that the process enters it if a man has no son, and that the state cannot be left. This state represents families whose male line has died out. Find the matrix of transition probabilities and find the probability that an unskilled laborer has a grandson who is a professional man. [*Ans.* .24.]

12. Explain why the transition probabilities given in Example 3 are correct.

SUGGESTED READING

Cramer, Harold, *The Elements of Probability Theory*, Wiley, New York, 1955, Part I.

Feller, W., *An Introduction to Probability Theory and its Applications*, Wiley, New York, 1950.

Neyman, J., *First Course in Probability and Statistics*, Holt, New York, 1950.

Whitworth, W. A., *Choice and Chance, with 1000 Exercises*, Stechert, New York, 1934.

Chapter V

VECTORS AND MATRICES

1. COLUMN AND ROW VECTORS

A *column vector* is an ordered collection of numbers written in a column. Examples of such vectors are

$$\begin{pmatrix} 1 \\ -2 \end{pmatrix}, \quad \begin{pmatrix} 0.6 \\ 0.4 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ -4 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ -1 \\ 2 \\ 4 \end{pmatrix}.$$

The individual numbers in these vectors are called *components*, and the number of components a vector has is one of its distinguishing characteristics. Thus the first two vectors above have two components; the next two have three components; and the last has four components. When talking more generally about n-component column vectors we shall write

$$u = \begin{pmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_n \end{pmatrix}.$$

Analogously, a *row vector* is an ordered collection of numbers written in a row. Examples of row vectors are

$$(1,0), \quad (-2,1), \quad (2,-3,4,0), \quad (-1,2,-3,4,-5).$$

Each number appearing in the vector is again called a *component* of the vector, and the number of components a row vector has is again one of its important characteristics. Thus, the first two examples are two-component, the third a four-component, and the fourth a five-component vector. The vector $v = (v_1,v_2, \ldots, v_n)$ is an n-component row vector.

Two row vectors, or two column vectors, are said to be *equal* if and only if corresponding components of the vector are equal. Thus for the vectors

$$u = (1,2), \quad v = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \quad w = (1,2), \quad x = (2,1),$$

we see that $u = w$ but $u \neq v$, and $u \neq x$.

If u and v are three-component column vectors, we shall define their sum $u + v$ by component-wise addition as follows:

$$u + v = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_1 + v_1 \\ u_2 + v_2 \\ u_3 + v_3 \end{pmatrix}.$$

Similarly, if u and v are three-component row vectors, their sum is defined to be

$$\begin{aligned} u + v &= (u_1,u_2,u_3) + (v_1,v_2,v_3) \\ &= (u_1+v_1,u_2+v_2,u_3+v_3). \end{aligned}$$

Note that the sum of two three-component vectors yields another three-component vector. For example,

$$\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix},$$

and

$$(4,-7,12) + (3,14,-14) = (7,7,-2).$$

The sum of two n-component vectors (either row or column) is defined by component-wise addition in an analogous manner, and yields another n-component vector. Observe that we do not define the addition of vectors unless they are both row or both column vectors, having the same number of components.

Because the order in which two numbers are added is immaterial

as far as the answer goes, it is also true that the order in which vectors are added does not matter; that is,

$$u + v = v + u$$

where u and v are both row or both column vectors. This is the so-called *commutative law of addition.* A numerical example is

$$\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \\ -1 \end{pmatrix} + \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}.$$

Once we have the definition of the addition of two vectors we can easily see how to add three or more vectors by grouping them in pairs as in the addition of numbers. For example,

$$\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 2 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$
$$= \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 3 \end{pmatrix} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix},$$

and

$$(1,0,0) + (0,2,0) + (0,0,3) = (1,2,0) + (0,0,3) = (1,2,3)$$
$$= (1,0,0) + (0,2,3) = (1,2,3).$$

In general, the sum of any number of vectors (row or column), each having the same number of components, is the vector whose first component is the sum of the first components of the vectors, whose second component is the sum of the second components, etc.

The multiplication of a number a times a vector v is defined by component-wise multiplication of a times the components of v. For the three-component case we have

$$au = a\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} au_1 \\ au_2 \\ au_3 \end{pmatrix}$$

for column vectors and

$$av = a(v_1,v_2,v_3) = (av_1,av_2,av_3)$$

for row vectors. If u is an n-component vector (row or column), then au is defined similarly by component-wise multiplication.

If u is any vector we define its negative $-u$ to be the vector $-u = (-1)u$. Thus in the three-component case for row vectors we have

$$-u = (-1)(u_1,u_2,u_3) = (-u_1,-u_2,-u_3).$$

Once we have the negative of a vector it is easy to see how to subtract vectors, i.e., we simply add "algebraically." For the three-component column vector case we have

$$u - v = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} - \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} u_1 - v_1 \\ u_2 - v_2 \\ u_3 - v_3 \end{pmatrix}.$$

Specific examples of subtraction of vectors occur in the exercises at the end of this section.

An important vector is the zero vector all of whose components are zero. For example, three-component zero vectors are

$$0 = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad 0 = (0,0,0).$$

When there is no danger of confusion we shall use the symbol 0, as above, to denote the zero (row or column) vector. The meaning will be clear from the context. The zero vector has the important property that, if u is any vector, then $u + 0 = u$. A proof for the three-component column vector case is as follows:

$$u + 0 = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} u_1 + 0 \\ u_2 + 0 \\ u_3 + 0 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = u.$$

One of the chief advantages of the vector notation is that one can denote a whole collection of numbers by a single letter such as u, v, . . . , and treat such a collection as if it were a single quantity. By using the vector notation it is possible to state very complicated relationships in a simple manner. The student will see many examples of this in the remainder of the present chapter and the two succeeding chapters.

EXERCISES

1. Compute the quantities below for the vectors

$$u = \begin{pmatrix} 3 \\ 1 \\ 2 \end{pmatrix}, \quad v = \begin{pmatrix} -2 \\ 3 \\ 0 \end{pmatrix}, \quad w = \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix}.$$

(a) $2u$. [*Ans.* $\begin{pmatrix} 6 \\ 2 \\ 4 \end{pmatrix}$.]

(b) $-v$.

(c) $2u - v$.

(d) $v + w$. [*Ans.* $\begin{pmatrix} -3 \\ 2 \\ 1 \end{pmatrix}$.]

(e) $u + v - w$.

(f) $2u - 3v - w$.

(g) $3u - v + 2w$. [*Ans.* $\begin{pmatrix} 9 \\ -2 \\ 8 \end{pmatrix}$.]

2. Compute (a) through (g) of Exercise 1 if the vectors u, v, and w are

$$u = (7,0,-3), \quad v = (2,1,-5), \quad w = (1,-1,0).$$

3. (a) Show that the zero vector is not changed when multiplied by any number.

(b) If u is any vector, show that $0 + u = u$.

4. If u and v are two row or two column vectors having the same number of components, prove that $u + 0v = u$ and $0u + v = v$.

5. If $2u - v = 0$, what is the relationship between the components of u and those of v? [*Ans.* $v_i = 2u_i$.]

6. Answer the question in Exercise 5 for the equation $-3u + 5v + u - 7v = 0$. Do the same for the equation $20v - 3u + 5v + 8u = 0$.

7. When possible compute the following sums; when not possible give reasons.

(a) $\begin{pmatrix} -1 \\ 3 \end{pmatrix} + \begin{pmatrix} 6 \\ -2 \\ 5 \\ -4 \end{pmatrix} = ?$

(b) $(2,-1,-1) + 0(4,7,-2) = ?$

(c) $(5,6) + 7 - 21 + \begin{pmatrix} 0 \\ 1 \end{pmatrix} = ?$

(d) $1\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} + 2\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + 3\begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix} = ?$

8. If $\begin{pmatrix} 1 \\ 1 \\ 2 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}$, find u_1, u_2, and u_3. [*Ans.* 0; −2; −2.]

9. If $2\begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 1 \\ 3 \end{pmatrix}$, find the components of v.

10. If $\begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, what can be said concerning the components u_1, u_2, u_3?

11. If $0 \cdot \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$, what can be said concerning the components u_1, u_2, u_3?

12. Suppose that we associate with each person a three-component row vector having the following entries: age, height, and weight. Would it make sense to add together the vectors associated with two different persons? Would it make sense to multiply one of these vectors by a constant?

13. Suppose that we associate with each person leaving a supermarket a row vector whose components give the quantities of each available item that he has purchased. Answer the same questions as those in Exercise 12.

14. Let us associate with each supermarket a column vector whose entries give the prices of each item in the store. Would it make sense to add together the vectors associated with two different supermarkets? Would it make sense to multiply one of these vectors by a constant? Discuss the differences in the situations given in Exercises 12, 13, and 14.

2. THE PRODUCT OF VECTORS; EXAMPLES

The reader may have wondered why it was necessary to introduce both column and row vectors when their properties are so similar. This question can be answered in several different ways. In the first place, in many applications there are two kinds of quantities which are studied simultaneously, and it is convenient to represent one of

them as a row vector and the other as a column vector. Second, there is a way of combining row and column vectors that is very useful for certain types of calculations. To bring out these points let us look at the following simple economic example.

Example 1. Suppose a man named Smith goes into a grocery store to buy a dozen each of eggs and oranges, a half dozen each of apples and pears, and three lemons. Let us represent his purchases by means of the following row vector:

$$\begin{aligned} x &= [6 \text{ (apples)}, 12 \text{ (eggs)}, 3 \text{ (lemons)}, 12 \text{ (oranges)}, 6 \text{ (pears)}] \\ &= (6,12,3,12,6). \end{aligned}$$

Suppose that apples are 4 cents each, eggs are 6 cents each, lemons are 9 cents each, oranges are 5 cents each, and pears are 7 cents each. We can then represent the prices of these items as a column vector

$$y = \begin{pmatrix} 4 \\ 6 \\ 9 \\ 5 \\ 7 \end{pmatrix} \begin{matrix} \text{cents per apple} \\ \text{cents per egg} \\ \text{cents per lemon} \\ \text{cents per orange} \\ \text{cents per pear.} \end{matrix}$$

The obvious question to ask now is, what is the total amount that Smith must pay for his purchases? What we would like to do is to multiply the quantity vector x by the price vector y, and we would like the result to be Smith's bill. We see that our multiplication should have the following form:

$$\begin{aligned} x \cdot y &= (6,12,3,12,6)\begin{pmatrix} 4 \\ 6 \\ 9 \\ 5 \\ 7 \end{pmatrix} \\ &= 6 \cdot 4 + 12 \cdot 6 + 3 \cdot 9 + 12 \cdot 5 + 6 \cdot 7 \\ &= 24 + 72 + 27 + 60 + 42 \\ &= 225 \text{ cents or } \$2.25. \end{aligned}$$

This is, of course, the computation that the cashier performs in figuring Smith's bill.

We shall adopt in general the above definition of multiplication of row times column vectors.

DEFINITION. Let u be a row vector and v a column vector each having the same number n of components; then we shall define the product $u \cdot v$ to be

$$u \cdot v = u_1 v_1 + u_2 v_2 + \ldots + u_n v_n.$$

Notice that we always write the row vector first and the column vector second, and this is the only kind of vector multiplication that we consider. Some examples of vector multiplication are given below:

$$(2,1,-1) \cdot \begin{pmatrix} 3 \\ -1 \\ 4 \end{pmatrix} = 2 \cdot 3 + 1 \cdot (-1) + (-1) \cdot 4 = 1.$$

$$(1,0) \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 1 \cdot 0 + 0 \cdot 1 = 0 + 0 = 0.$$

Note that the result of vector multiplication is always a *number*.

Example 2. Consider an oversimplified economy which has three industries which we call coal, electricity, and steel, and three consumers 1, 2, and 3. Suppose that each consumer uses some of the output of each industry and also that each industry uses some of the output of each other industry. We assume that the amounts used are positive or zero, since using a negative quantity has no immediate interpretation. We can represent the needs of each consumer and industry by a three-component demand (row) vector, the first component measuring the amount of coal needed by the consumer or industry; the second component the amount of electricity needed; and the third component the amount of steel needed, in some convenient units. For example, the demand vectors of the three consumers might be

$$d_1 = (3,2,5), \quad d_2 = (0,17,1), \quad d_3 = (4,6,12);$$

and the demand vectors of each of the industries might be

$$d_C = (0,1,4), \quad d_E = (20,0,8), \quad d_S = (30,5,0),$$

where the subscript C stands for coal; the subscript E, for electricity;

and the subscript S, for steel. Then the total demand for these goods by the consumers is given by the sum

$$d_1 + d_2 + d_3 = (3,2,5) + (0,17,1) + (4,6,12) = (7,25,18).$$

Also, the total industrial demand for these goods is given by the sum

$$d_C + d_E + d_S = (0,1,4) + (20,0,8) + (30,5,0) = (50,6,12).$$

Therefore the total over-all demand is given by the sum

$$(7,25,18) + (50,6,12) = (57,31,30).$$

Suppose now that the price of coal is \$1 per unit, the price of electricity is \$2 per unit, and the price of steel is \$4 per unit. Then these prices can be represented by the column vector

$$p = \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix}.$$

Consider the steel industry: it sells a total of 30 units of steel at \$4 per unit so that its total income is \$120. Its bill for the various goods is given by the vector product

$$d_S \cdot p = (30,5,0) \cdot \begin{pmatrix} 1 \\ 2 \\ 4 \end{pmatrix} = 30 + 10 = \$40.$$

Hence the profit of the steel industry is \$120 − \$40 = \$80. In the exercises below the profits of the other industries will be found.

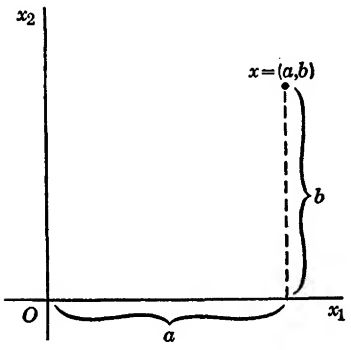


Figure 1

This model of an economy is unrealistic in two senses. First we have not chosen realistic numbers for the various quantities involved. Second, and more important, we have neglected the fact that the more an industry produces the more inputs it requires. The latter complication will be introduced in Chapter VII.

Example 3. Consider the rectangular coordinate system in the

plane shown in Figure 1. A two-component row vector $x = (a,b)$ can be regarded as a point in the plane located by means of the coordinate axes as shown. The point x can be found by starting at the origin of coordinates O and moving a distance a along the x_1 axis; then

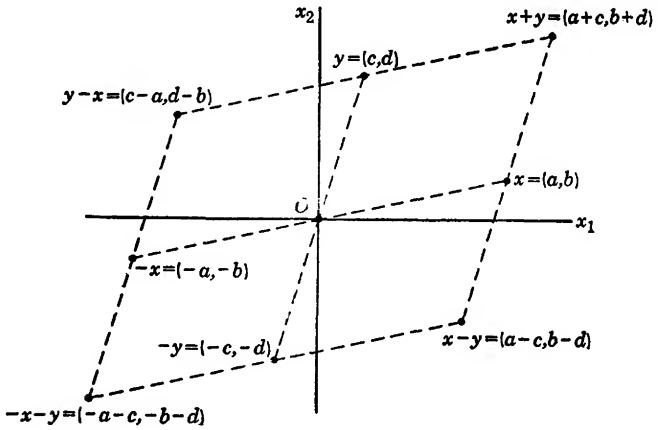


Figure 2

moving a distance b along a line parallel to the x_2 axis. If we have two such points, say $x = (a,b)$ and $y = (c,d)$, then the points $x + y$, $-x$, $-y$, $x - y$, $y - x$, $-x - y$ have the geometric significance shown in Figure 2.

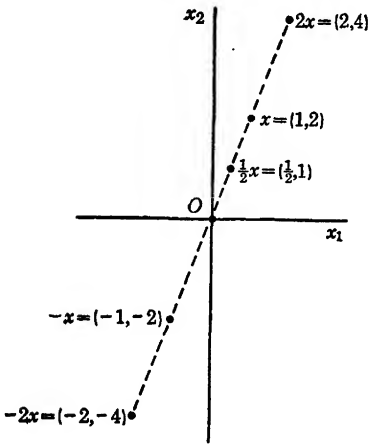


Figure 3

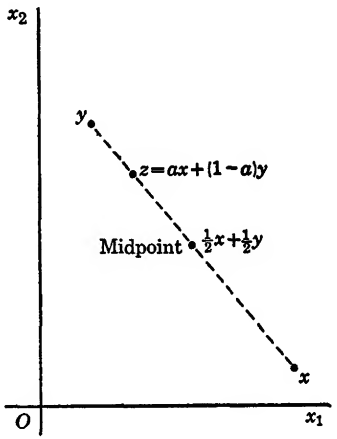


Figure 4

The idea of multiplying a row vector by a number can also be given a geometric meaning, see Figure 3. There we have plotted the point corresponding to the vector $x = (1,2)$ and $2x$, $\frac{1}{2}x$, $-x$, and $-2x$. Observe that all these points lie on a line through the origin of coordinates. Another vector quantity which has geometrical significance is the vector $z = ax + (1 - a)y$, where a is any number between 0 and 1. Observe in Figure 4 that the points z all lie on the line segment between the points x and y. If $a = \frac{1}{2}$ the corresponding point on the line segment is the mid-point of the segment. Thus, if $x = (a,b)$ and $y = (c,d)$ then the point

$$\begin{aligned} \tfrac{1}{2}x + \tfrac{1}{2}y &= \tfrac{1}{2}(a,b) + \tfrac{1}{2}(c,d) \\ &= \left(\frac{a+c}{2}, \frac{b+d}{2}\right) \end{aligned}$$

is the mid-point of the line segment between x and y.

EXERCISES

1. Compute the quantities below for the following vectors:

$$u = (1,-1,4), \qquad x = (0,1,2),$$

$$v = \begin{pmatrix} 5 \\ 0 \\ 1 \end{pmatrix}, \qquad y = \begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}.$$

(a) $u \cdot v + x \cdot y = ?$ [*Ans.* 12.]

(b) $(-u + 5x) \cdot (3v - 2y) = ?$

(c) $5u \cdot v + 10[x \cdot (2v - y)] = ?$ [*Ans.* 55.]

(d) $2[(u - x) \cdot (v + y)] = ?$

2. Plot the points corresponding to the row vectors $x = (3,4)$ and $y = (-2,7)$. Then compute and plot the following vectors.

(a) $\frac{1}{2}x + \frac{1}{2}y$.

(b) $x + y$.

(c) $x - 2y$.

(d) $\frac{7}{8}x + \frac{1}{8}y$.

(e) $3x - 2y$.

(f) $4y - 3x$.

3. If $x = (1,-1,2)$ and $y = (0,1,3)$ are points in space, what is the mid-point of the line segment joining x to y? [*Ans.* $(\frac{1}{2},0,\frac{5}{2})$.]

4. If u is a three-component row vector and v is a three-component column

vector having the same number of components, and a is a number, prove that $a(u \cdot v) = (au) \cdot v = u \cdot (av)$.

5. Suppose that Brown, Jones, and Smith go to the grocery store and purchase the following items:

Brown: two apples, six lemons, and five pears;

Jones: two dozen eggs, two lemons, and two dozen oranges;

Smith: ten apples, one dozen eggs, two dozen oranges, and a half dozen pears.

(a) How many different kinds of items did they purchase? [*Ans.* 5.]

(b) Write each of their purchases as row vectors with as many components as the answer found in (a).

(c) Using the price vector given in Example 1, compute each man's grocery bill. [*Ans.* \$0.97; \$2.82; \$2.74.]

(d) By means of vector addition find the total amount of their purchases as a row vector.

(e) Compute in two different ways the total amount spent by the three men at the grocery store. [*Ans.* \$6.53.]

6. Prove that vector multiplication satisfies the following two properties:

$$\text{(i)} \quad u \cdot (av) = a(u \cdot v)$$

$$\text{(ii)} \quad u \cdot (v + w) = u \cdot v + u \cdot w$$

where u is a three-component row vector, v and w are three-component column vectors, and a is a number.

7. The production of a book involves several steps: first it must be set in type, then it must be printed, and finally it must be supplied with covers and bound. Suppose that the typesetter charges \$6 an hour, paper costs $\frac{1}{4}$ cent per sheet, that the printer charges 11 cents for each minute that his press runs, that the cover costs 28 cents, and that the binder charges 15 cents to bind each book. Suppose now that a publisher wishes to print a book that requires 300 hours of work by the typesetter, 220 sheets of paper per book, and 5 minutes of press time per book.

(a) Write a five-component row vector which gives the requirements for the first book. Write another row vector which gives the requirements for the second, third, . . . copies of the book. Write a five-component column vector whose components give the prices of the various requirements for each book, in the same order as they are listed in the requirement vectors above.

(b) Using vector multiplication, find the cost of publishing one copy of a book. [*Ans.* \$1,801.53.]

(c) Using vector addition and multiplication, find the cost of printing a first edition run of 5000 copies. [*Ans.* \$9,450.]

(d) Assuming that the type plates from the first edition are used again, find the cost of printing a second edition of 5000 copies. [*Ans.* \$7,650.]

8. Perform the following calculations for Example 2.

(a) Compute the amount that each industry and each consumer has to pay for the goods it receives.

(b) Compute the profit made by each of the industries.

(c) Find the total amount of money that is paid out by all the industries and consumers.

(d) Find the proportion of the total amount of money found in (c) paid out by the industries. Find the proportion of the total money that is paid out by the consumers.

9. A building contractor has accepted orders for five ranch style houses, seven Cape Cod houses, and twelve Colonial style houses. Write a three-component row vector x whose components give the numbers of each type of house to be built. Suppose that he knows that a ranch style house requires 20 units of wood; a Cape Cod, 18 units; and a Colonial style, 25 units of wood. Write a column vector u whose components give the various quantities of wood needed for each type of house. Find the total amount of wood needed by computing the matrix product xu. [*Ans.* 526.]

10. Let $x = (x_1,x_2)$ and let a and b be the vectors

$$a = \begin{pmatrix} 3 \\ 4 \end{pmatrix}, \qquad b = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

If $x \cdot a = -1$ and $x \cdot b = 7$ determine x_1 and x_2. [*Ans.* $x_1 = -31$; $x_2 = 23$.]

11. Let $x = (x_1,x_2)$ and let a and b be the vectors

$$a = \begin{pmatrix} 4 \\ 8 \end{pmatrix}, \qquad b = \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

If $x \cdot a = x_1$ and $x \cdot b = x_2$, determine x_1 and x_2.

3. MATRICES AND THEIR COMBINATION WITH VECTORS

A matrix is a rectangular array of numbers written in the form

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

Here the letters a_{ij} stand for real numbers and m and n are integers. Observe that m is the number of rows and n is the number of columns

of the matrix. For this reason we call it an $m \times n$ matrix. If $m = n$ the matrix is *square*. The following are examples of matrices.

$$(1,2,3), \quad \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \quad \begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix},$$

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 7 & -8 & 9 & 10 \\ 3 & -1 & 14 & 2 & -6 \\ 0 & 3 & -5 & 7 & 0 \end{pmatrix}.$$

The first example is a row vector which is a 1×3 matrix; the second is a column vector which is a 3×1 matrix; the third example is a 2×2 square matrix; the fourth is a 4×4 square matrix; and the last is a 3×5 matrix.

Two matrices having the same shape (i.e., having the same number of rows and columns) are said to be equal if and only if the corresponding entries are equal.

Recall that in Chapter IV, Section 13, we found that a matrix arose naturally in the consideration of a Markov chain process. To give another example of how matrices occur in practice and are used in connection with vectors we consider the following example.

Example 1. Suppose that a building contractor has accepted orders for five ranch style houses, seven Cape Cod houses, and twelve Colonial style houses. We can represent his orders by means of a row vector $x = (5,7,12)$. The contractor is familiar of course, with the kinds of "raw materials" that go into each type of house. Let us suppose that these raw materials are steel, wood, glass, paint, and labor. The numbers in the matrix below give the amounts of each raw material going into each type of house, expressed in convenient units. (The numbers are put in arbitrarily, and are not meant to be realistic.)

	Steel	Wood	Glass	Paint	Labor
Ranch:	5	20	16	7	17
Cape Cod:	7	18	12	9	21
Colonial:	6	25	8	5	13

$= R$

Observe that each row of the matrix is a five-component row vector which gives the amounts of each raw material needed for a given kind

of house. Similarly, each column of the matrix is a three-component column vector which gives the amounts of a given raw material needed for each kind of house. Clearly, a matrix is a very succinct way of summarizing this information.

Suppose now that the contractor wishes to compute how much of each raw material to obtain in order to fulfill his contracts. Let us denote the matrix above by R; then he would like to obtain something like the product xR, and he would like the product to tell him what orders to make out. The product should have the following form:

$$\begin{aligned} xR &= (5,7,12)\begin{pmatrix} 5 & 20 & 16 & 7 & 17 \\ 7 & 18 & 12 & 9 & 21 \\ 6 & 25 & 8 & 5 & 13 \end{pmatrix} \\ &= (5\cdot 5 + 7\cdot 7 + 12\cdot 6,\ 5\cdot 20 + 7\cdot 18 + 12\cdot 25, \\ &\qquad 5\cdot 16 + 7\cdot 12 + 12\cdot 8,\ 5\cdot 7 + 7\cdot 9 + 12\cdot 5, \\ &\qquad 5\cdot 17 + 7\cdot 21 + 12\cdot 13) \\ &= (146,526,260,158,388). \end{aligned}$$

Thus we see that the contractor should order 146 units of steel, 526 units of wood, 260 units of glass, 158 units of paint, and 388 units of labor. Observe that the answer we get is a five-component row vector and that each entry in this vector is obtained by taking the vector product of x times the corresponding column of the matrix R.

The contractor is also interested in the prices that he will have to pay for these materials. Suppose that steel costs \$15 per unit, wood costs \$8 per unit, glass costs \$5 per unit, paint costs \$1 per unit, and labor costs \$10 per unit. Then we can write the cost as a column vector as follows:

$$y = \begin{pmatrix} 15 \\ 8 \\ 5 \\ 1 \\ 10 \end{pmatrix}.$$

Here the product Ry should give the costs of each type of house, so that the multiplication should have the form

$$Ry = \begin{pmatrix} 5 & 20 & 16 & 7 & 17 \\ 7 & 18 & 12 & 9 & 21 \\ 6 & 25 & 8 & 5 & 13 \end{pmatrix} \begin{pmatrix} 15 \\ 8 \\ 5 \\ 1 \\ 10 \end{pmatrix}$$

$$= \begin{pmatrix} 5\cdot 15 + 20\cdot 8 + 16\cdot 5 + 7\cdot 1 + 17\cdot 10 \\ 7\cdot 15 + 18\cdot 8 + 12\cdot 5 + 9\cdot 1 + 21\cdot 10 \\ 6\cdot 15 + 25\cdot 8 + 8\cdot 5 + 5\cdot 1 + 13\cdot 10 \end{pmatrix}$$

$$= \begin{pmatrix} 492 \\ 528 \\ 465 \end{pmatrix}.$$

Thus the cost of materials for the ranch style house is \$492, for the Cape Cod house is \$528, and for the Colonial house \$465.

The final question which the contractor might ask is what is the total cost of raw materials for all the houses he will build. It is easy to see that this is given by the vector xRy. We can find it in two ways as shown below.

$$xRy = (xR)y = (146,526,260,158,388)\cdot \begin{pmatrix} 15 \\ 8 \\ 5 \\ 1 \\ 10 \end{pmatrix} = 11,736$$

$$xRy = x(Ry) = (5,7,12)\cdot \begin{pmatrix} 492 \\ 528 \\ 465 \end{pmatrix} = 11,736.$$

The total cost is then \$11,736.

We shall adopt, in general, the above definitions for the multiplication of a matrix times a row or a column vector.

DEFINITION. Let A be an $m \times n$ matrix, let x be an m-component row vector, and let u be a n-component column vector; then we define the products xA and Au as follows:

$$xA = (x_1,x_2,\ldots,x_m)\begin{pmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ & & \ldots & \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{pmatrix}$$

$$= (x_1a_{11} + x_2a_{21} + \ldots + x_ma_{m1},\ x_1a_{12} + x_2a_{22} + \ldots + x_ma_{m2},\ \ldots,$$
$$x_1a_{1n} + x_2a_{2n} + \ldots + x_ma_{mn});$$

$$Au = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ & & \dots & \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ . \\ . \\ . \\ u_n \end{pmatrix} = \begin{pmatrix} a_{11}u_1 + a_{12}u_2 + \dots + a_{1n}u_n \\ a_{21}u_1 + a_{22}u_2 + \dots + a_{2n}u_n \\ \dots \\ a_{m1}u_1 + a_{m2}u_2 + \dots + a_{mn}u_n \end{pmatrix}.$$

The reader will find these formulas easy to work with if he observes that each entry in the products xA or Au is obtained by vector multiplication of x or u by a column or row of the matrix A. Notice that in order to multiply a row vector times a matrix, the number of rows of the matrix must equal the number of components of the vector, and the result is another row vector; similarly, to multiply a matrix times a column vector, the number of columns of the matrix must equal the number of components of the vector, and the result of such a multiplication is another column vector.

Some numerical examples of the multiplication of vectors and matrices are:

$$(1,0,-1)\begin{pmatrix} 3 & 1 \\ 2 & 3 \\ 2 & 8 \end{pmatrix} = (1\cdot 3 + 0\cdot 2 - 1\cdot 2,\ 1\cdot 1 + 0\cdot 3 - 1\cdot 8)$$
$$= (1,-7);$$

$$\begin{pmatrix} 3 & 1 & 2 \\ 2 & 3 & 8 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} 3 - 1 + 4 \\ 2 - 3 + 16 \end{pmatrix} = \begin{pmatrix} 6 \\ 15 \end{pmatrix};$$

$$\begin{pmatrix} 3 & 2 & -1 \\ 1 & 0 & 2 \\ 0 & 3 & 1 \\ 5 & -4 & 7 \\ -3 & 2 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 5 \\ -3 \\ -2 \\ -9 \\ -1 \end{pmatrix}.$$

Observe that if x is an m-component row vector and A is $m \times n$, then xA is an n-component row vector; similarly, if u is an n-component column vector, then Au is an m-component column vector. These facts can be observed in the examples above.

Example 2. In Exercise 6 of Chapter IV, Section 13, we considered a Markov chain with transition matrix

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

The initial state was chosen by a random device that selected states a_1 and a_2 each with probability $\frac{1}{2}$. Let us indicate the choice of initial state by the vector $p^{(0)} = (\frac{1}{2},\frac{1}{2})$ where the first component gives the probability of choosing state a_1 and the second the probability of choosing state a_2. Let us compute the product $p^{(0)}P$. We have

$$\begin{aligned} p^{(0)}P &= (\tfrac{1}{2},\tfrac{1}{2})\begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} \\ &= (\tfrac{1}{6} + \tfrac{1}{4},\ \tfrac{1}{3} + \tfrac{1}{4}) \\ &= (\tfrac{5}{12},\tfrac{7}{12}). \end{aligned}$$

Using the methods of Chapter IV, one can show that after one step there is probability $\frac{5}{12}$ that the process will be in state a_1 and probability $\frac{7}{12}$ that it will be in state a_2. Let $p^{(1)}$ be the vector whose first component gives the probability of the process being in state a_1 after one step and whose second component gives the probability of it being state a_2 after one step. In our example we have $p^{(1)} = (\frac{5}{12},\frac{7}{12}) = p^{(0)}P$.

In general the formula $p^{(1)} = p^{(0)}P$ holds for any Markov process with transition matrix P and initial probability vector $p^{(0)}$.

EXERCISES

1. Perform the following multiplications:

(a) $\begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix}\begin{pmatrix} 7 \\ 2 \end{pmatrix} = ?$

(b) $(3,-4)\begin{pmatrix} 1 & -1 \\ -2 & 2 \end{pmatrix} = ?$ [*Ans.* (11,−11).]

(c) $\begin{pmatrix} 1 & 3 & 0 \\ 7 & -1 & 3 \\ -8 & 14 & -5 \\ 9 & 2 & 7 \\ 10 & -6 & 0 \end{pmatrix}\cdot\begin{pmatrix} 3 \\ -1 \\ 1 \end{pmatrix} = ?$

(d) $(2,2)\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = ?$ [*Ans.* (0,0).]

(e) $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 5 \\ 5 \end{pmatrix} = ?$

(f) $(0,2,-3)\begin{pmatrix} 1 & 7 & -8 & 9 & 10 \\ 3 & -1 & 14 & 2 & -6 \\ 0 & 3 & -5 & 7 & 0 \end{pmatrix} = ?$

(g) $(x_1,x_2)\begin{pmatrix} a & b \\ c & d \end{pmatrix} = ?$ [*Ans.* $(ax_1 + cx_2,\ bx_1 + dx_2)$.]

(h) $\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = ?$

(i) $\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = ?$

(j) $(x_1,x_2,x_3)\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = ?$

2. What number does the matrix in parts (i) and (j) above resemble?

3. Notice that in Exercise 1(d) above the product of a row vector, none of whose components is zero, times a matrix, none of whose components is zero, yields the zero row vector. Find another example which is similar to this one. Answer the analogous question for Exercise 1(e).

4. When possible, solve for the indicated quantities.

(a) $(x_1,x_2)\begin{pmatrix} 0 & -1 \\ 7 & 3 \end{pmatrix} = (7,0)$. Find the vector x. [*Ans.* (3,1).]

(b) $(2,-1)\begin{pmatrix} a & b \\ c & d \end{pmatrix} = (6,3)$. Find the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$. In this case can you find more than one solution?

(c) $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$. Find the vector u.

(d) $\begin{pmatrix} -1 & 4 \\ 2 & -8 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 3 \\ -6 \end{pmatrix}$. Find u.

How many solutions can you find?

[*Ans.* $u = \begin{pmatrix} 4k - 3 \\ k \end{pmatrix}$, for any number k.]

5. Solve for the indicated quantities below and give an interpretation for each.

(a) $(1,-1)\begin{pmatrix} 0 & 2 \\ -2 & 4 \end{pmatrix} = a(1,-1)$; find a. [*Ans.* $a = 2$.]

(b) $\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = 5\begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$; find u. How many answers can you find?

[*Ans.* $u = \begin{pmatrix} k \\ 2k \end{pmatrix}$ for any number k.]

(c) $\begin{pmatrix} \frac{5}{8} & \frac{1}{8} \\ \frac{3}{8} & \frac{7}{8} \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$; find u. How many answers are there?

6. In Exercise 5 of the preceding section construct the 3×5 matrix whose rows give the various purchases of Brown, Jones, and Smith. Multiply on the right by the five-component price (column) vector to find the three-component column vector whose entries give each person's grocery bill. Multiply on the left by the row vector $x = (1,1,1)$ and on the right by the price vector to find the total amount that they spent in the store.

7. In Example 1 of this section, assume that the contractor is to build seven ranch style, three Cape Cod, and five Colonial type houses. Recompute, using matrix multiplication, the total cost of raw materials, in two different ways as in the example.

8. In Example 2 of this section, assume that the initial probability vector is $p^{(0)} = (\frac{1}{3},\frac{2}{3})$. Find the vector $p^{(1)}$. [*Ans.* $(\frac{4}{9}, \frac{5}{9})$.]

9. For the Markov process whose transition matrix is

$$P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}$$

assume the initial probability vector is $p^{(0)} = (\frac{1}{2},\frac{1}{6},\frac{1}{3})$. Draw the tree of the process and find the tree measures. Compute $p^{(1)}$ by means of the tree measure and also from the formula $p^{(1)} = p^{(0)}P$ and show that the two answers agree.

10. Consider the Markov process with two states whose transition matrix is

$$P = \begin{pmatrix} a & 1-a \\ 1-b & b \end{pmatrix}$$

where a and b are nonnegative numbers. Suppose the initial probability vector for the process is

$$p^{(0)} = (p_1^{(0)},p_2^{(0)})$$

where $p_1^{(0)}$ is the initial probability of choosing state 1 and $p_2^{(0)}$ is the initial probability of choosing state 2. Derive the formulas for the components of the vector $p^{(1)}$. [*Ans.* $p^{(1)} = \{ap_1^{(0)} + (1-b)p_2^{(0)}, (1-a)p_1^{(0)} + bp_2^{(0)}\}$.]

11. In Example 2 use tree measures to show that $p^{(2)} = p^{(1)}P$.

12. The following matrix gives the vitamin contents of three food items, in conveniently chosen units:

Vitamin:	A	B	C	D
Food I:	.5	.5	0	0
Food II:	.3	0	.2	.1
Food III:	.1	.1	.2	.5

If we eat 5 units of food I, 10 units of food II, and 8 units of food III, how much of each type of vitamin have we consumed? If we pay only for the

vitamin content of each food, paying 10 cents, 20 cents, 25 cents, and 50 cents, respectively, for units of the four vitamins, how much does a unit of each type of food cost? Compute in two ways the total cost of the food we ate.

[*Ans.* (6.3,3.3,3.6,5.0); $\begin{pmatrix} 15 \\ 13 \\ 33 \end{pmatrix}$; \$4.69.]

4. THE ADDITION AND MULTIPLICATION OF MATRICES

Two matrices of the same shape, that is, having the same number of rows and columns, can be added together by adding corresponding components. For example, if A and B are two 2×3 matrices, we have

$$A + B = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} + \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \end{pmatrix}$$
$$= \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & a_{13} + b_{13} \\ a_{21} + b_{21} & a_{22} + b_{22} & a_{23} + b_{23} \end{pmatrix}.$$

Observe that the addition of vectors (row or column) is simply a special case of the addition of matrices. Numerical examples of the addition of matrices are the following:

$$(1,0,-2) + (0,5,0) = (1,5,-2);$$

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix};$$

$$\begin{pmatrix} 7 & 0 & 0 \\ -3 & 1 & -6 \\ 4 & 0 & 7 \\ 0 & -2 & -2 \\ 1 & 1 & 1 \end{pmatrix} + \begin{pmatrix} -8 & 0 & 1 \\ 4 & 5 & -1 \\ 0 & 3 & 0 \\ -1 & 1 & -1 \\ 0 & -4 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 0 & 1 \\ 1 & 6 & -7 \\ 4 & 3 & 7 \\ -1 & -1 & -3 \\ 1 & -3 & 3 \end{pmatrix}.$$

Other examples occur in the exercises. The reader should observe that we do *not* add matrices of different shapes.

If A is a matrix and k is any number, we define the matrix kA as

$$kA = k\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = \begin{pmatrix} ka_{11} & ka_{12} & \cdots & ka_{1n} \\ ka_{21} & ka_{22} & \cdots & ka_{2n} \\ \cdot & \cdot & \cdots & \cdot \\ ka_{m1} & ka_{m2} & \cdots & ka_{mn} \end{pmatrix}.$$

Observe that this is merely component-wise multiplication as was the

analogous concept for vectors. Some examples of multiplication of matrices by constants are

$$-2\begin{pmatrix} 7 & -2 & 8 \\ 0 & 5 & -1 \end{pmatrix} = \begin{pmatrix} -14 & 4 & -16 \\ 0 & -10 & 2 \end{pmatrix};$$

$$6\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 3 & -4 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 6 \\ 18 & -24 \end{pmatrix}.$$

The multiplication of a vector by a number is, of course, a special case of the multiplication of a matrix by a number.

Under certain conditions two matrices can be multiplied together to give a new matrix. As an example, let A be a 2×3 matrix and B be a 3×2 matrix. Then the product AB is found as

$$AB = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}$$

$$= \begin{pmatrix} a_{11}b_{11} + a_{12}b_{21} + a_{13}b_{31} & a_{11}b_{12} + a_{12}b_{22} + a_{13}b_{32} \\ a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31} & a_{21}b_{12} + a_{22}b_{22} + a_{23}b_{32} \end{pmatrix}.$$

Observe that the product is a 2×2 matrix. Also notice that each entry in the new matrix is the product of one of the rows of A times one of the columns of B; for example, the entry in the second row and first column is found as the product

$$(a_{21} \quad a_{22} \quad a_{23})\begin{pmatrix} b_{11} \\ b_{21} \\ b_{31} \end{pmatrix} = a_{21}b_{11} + a_{22}b_{21} + a_{23}b_{31}.$$

The following definition holds for the general case of matrix multiplication:

DEFINITION. Let A be an $m \times k$ matrix and B be a $k \times n$ matrix; then the product matrix $C = AB$ is an $m \times n$ matrix whose components are

$$c_{ij} = (a_{i1} \quad a_{i2} \quad \ldots \quad a_{ik})\begin{pmatrix} b_{1j} \\ b_{2j} \\ \cdot \\ \cdot \\ \cdot \\ b_{kj} \end{pmatrix}$$

$$= a_{i1}b_{1j} + a_{i2}b_{2j} + \ldots + a_{ik}b_{kj}.$$

The important things to remember about this definition are: first, in order to be able to multiply matrix A times matrix B, the number of columns of A must be equal to the number of rows of B; second, the product matrix $C = AB$ has the same number of rows as A and the same number of columns as B; finally, to get the entry in the ith row and jth column of AB we multiply the ith row of A times the jth column of B. Notice that the product of a vector times a matrix is a special case of matrix multiplication.

Below are several examples of matrix multiplication:

$$\begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 7 & 0 \\ -2 & -3 \end{pmatrix} = \begin{pmatrix} 16 & 3 \\ -6 & -9 \end{pmatrix};$$

$$\begin{pmatrix} 3 & 0 & 1 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & -1 & 0 \\ 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 1 & 1 \\ -1 & -2 & 0 \\ 2 & 2 & 2 \end{pmatrix};$$

$$\begin{pmatrix} 3 & 1 & 4 \\ 2 & 0 & 5 \end{pmatrix}\begin{pmatrix} 1 & 3 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 10 & 4 & 4 \\ 2 & 6 & 5 & 5 \end{pmatrix}.$$

One obvious question that now arises is that of multiplying more than two matrices together. Let A be an $m \times h$ matrix, let B be an $h \times k$ matrix, and let C be a $k \times n$ matrix. Then we can certainly define the products $(AB)C$ and $A(BC)$. It turns out that these two products are equal, and we define the product ABC to be their common value, i.e.,

$$ABC = A(BC) = (AB)C.$$

The rule expressed in the above equation is called the *associative law* for multiplication. We shall not prove the associative law here although the student will be asked to check an example of it in Exercise 5.

If A and B are square matrices of the same size, then they can be multiplied in either order. It is not true, however, that the product AB is necessarily equal to the product BA. For example, if

$$A = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}$$

then we have

$$AB = \begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}$$

whereas

$$BA = \begin{pmatrix} 1 & 0 \\ 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix},$$

and it is clear that $AB \neq BA$.

EXERCISES

1. Perform the following operations.

(a) $2\begin{pmatrix} 6 & 1 \\ 0 & -3 \\ -1 & 2 \end{pmatrix} - 3\begin{pmatrix} 4 & 2 \\ 0 & 1 \\ -5 & -1 \end{pmatrix} = ?$ [*Ans.* $\begin{pmatrix} 0 & -4 \\ 0 & -9 \\ 13 & 7 \end{pmatrix}$.]

(b) $\begin{pmatrix} 6 & 1 & -1 \\ 1 & -3 & 2 \end{pmatrix} - 5\begin{pmatrix} 4 & 0 & -5 \\ 2 & 1 & -1 \end{pmatrix} = ?$

(c) $\begin{pmatrix} 6 & 1 \\ 0 & -3 \end{pmatrix}\begin{pmatrix} 4 & 0 & -4 \\ 2 & 1 & -1 \end{pmatrix} = ?$

(d) $\begin{pmatrix} 6 & 0 & -1 \\ 1 & -3 & 2 \end{pmatrix}\begin{pmatrix} 4 & 2 \\ 0 & 1 \\ -5 & -1 \end{pmatrix} = ?$ [*Ans.* $\begin{pmatrix} 29 & 13 \\ -6 & -3 \end{pmatrix}$.]

(e) $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} = ?$

(f) $\begin{pmatrix} 4 & 1 & 4 \\ -1 & -2 & -1 \\ 2 & -1 & -2 \end{pmatrix}\begin{pmatrix} 3 & 0 & 1 \\ -1 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix} = ?$ [*Ans.* $\begin{pmatrix} 11 & 2 & 12 \\ -1 & -4 & -3 \\ 7 & -2 & -2 \end{pmatrix}$.]

(g) $\begin{pmatrix} 1 & -2 \\ 0 & 0 \\ 7 & 5 \\ -4 & 8 \\ 0 & -2 \end{pmatrix}\begin{pmatrix} -7 & 9 & -5 & 6 & 0 \\ -1 & 0 & 3 & -4 & 1 \end{pmatrix} = ?$

2. Let A be any 3×3 matrix and let I be the matrix

$$I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Show that $AI = IA = A$. The matrix I acts for the products of matrices in the same way that the number 1 acts for products of numbers. For this reason it is called the identity matrix.

3. Let A be any 3×3 matrix and let 0 be the matrix

$$0 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Show that $A0 = 0A = 0$ for any A. Also show that $A + 0 = 0 + A = A$ for any A. The matrix 0 acts for matrices in the same way that the number 0 acts for numbers. For this reason it is called the zero matrix.

4. If $A = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$ and $B = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$ show that $AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$.

Thus the product of two matrices can be the zero matrix even though neither of the matrices is itself zero. Find another example that illustrates this point.

5. Verify the associative law for the special case when

$$A = \begin{pmatrix} -1 & 0 & 5 \\ 7 & -2 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 7 & 0 \\ -3 & -1 & 0 \\ 1 & 0 & 5 \end{pmatrix}, \quad C = \begin{pmatrix} -1 & -1 \\ 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

6. Consider the matrices

$$A = \begin{pmatrix} 1 & 0 & 1 \\ -1 & 17 & 57 \end{pmatrix}, \quad B = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \\ 0 & 0 & 0 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & 0 & -1 \\ 0 & -1 & 1 \\ -1 & 1 & 0 \end{pmatrix},$$

$$D = \begin{pmatrix} -1 & -1 \\ 2 & 2 \\ 1 & 1 \end{pmatrix}.$$

The shapes of these are 2×3, 4×3, 3×3, and 3×2, respectively. What is the shape of

(a) AC.
(b) DA.
(c) AD.
(d) BC.
(e) CB.
(f) DAC.
(g) $BCDA$. [*Ans.* 4×3.]

7. In Exercise 6 find:

(a) The component in the second row and second column of AC. [*Ans.* 40.]
(b) The component in the fourth row and first column of BC.
(c) The component in the last row and last column of DA. [*Ans.* 58.]
(d) The component in the first row and first column of CB.

8. If A is a square matrix, it can be multiplied by itself; hence we can define (using the associative law)

$$A^2 = A \cdot A$$
$$A^3 = A^2 \cdot A = A \cdot A \cdot A$$
$$\cdots$$
$$A^n = A^{n-1} \cdot A = A \cdot A \cdot \ldots A \qquad (n \text{ factors}).$$

These are naturally called "powers" of a matrix—the first one being called the square; the second, the cube; etc. Compute the indicated powers of the following matrices.

(a) If $A = \begin{pmatrix} 1 & 0 \\ 3 & 4 \end{pmatrix}$, find A^2, A^3, and A^4.

[*Ans.* $\begin{pmatrix} 1 & 0 \\ 15 & 16 \end{pmatrix}$; $\begin{pmatrix} 1 & 0 \\ 63 & 64 \end{pmatrix}$; $\begin{pmatrix} 1 & 0 \\ 255 & 256 \end{pmatrix}$.]

(b) If I and 0 are the matrices defined in Exercises 2 and 3, find I^2, I^3, I^n, 0^2, 0^3, and 0^n.

(c) If $A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 2 & -1 & 0 \end{pmatrix}$, find A^2, A^3, and A^n.

(d) If $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$, find A^n.

9. Cube the matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & 0 & \frac{2}{3} \end{pmatrix}.$$

Compare your answer with the matrix $P^{(3)}$ in Example 1, Chapter IV, Section 13, and comment on the result.

10. Consider a two-stage Markov process whose transition matrix is

$$P = \begin{pmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{pmatrix}.$$

(a) Assuming that the process starts in state 1, draw the tree and set up tree measures for three stages of the process. Do the same, assuming that the process starts in state 2.

(b) Using the trees drawn in (a), compute the quantities $p_{11}^{(3)}$, $p_{12}^{(3)}$, $p_{21}^{(3)}$, $p_{22}^{(3)}$. Write the matrix $P^{(3)}$.

(c) Compute the cube P^3 of the matrix P.

(d) Compare the answers you found in parts (b) and (c) and show that $P^{(3)} = P^3$.

11. Show that the fifth and all higher powers of the matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$

have all entries positive. Show that no smaller power has this property.

12. In Example 1 of Section 3 assume that the contractor wishes to take into account the cost of transporting raw materials to the building site as well as the purchasing cost. Suppose the costs are as given in the matrix below:

$$Q = \begin{array}{c} \begin{matrix} \text{Purchase} & \text{Transport} \end{matrix} \\ \begin{pmatrix} 15 & 4.5 \\ 8 & 2 \\ 5 & 3 \\ 1 & 0.5 \\ 10 & 0 \end{pmatrix} \end{array} \begin{matrix} \text{Steel} \\ \text{Wood} \\ \text{Glass} \\ \text{Paint} \\ \text{Labor} \end{matrix}$$

Referring to the example:

(a) By computing the product RQ find a 3×2 matrix whose entries give the purchase and transportation costs of the materials for each kind of house.

(b) Find the product xRQ, which is a two-component row vector whose first component gives the total purchase price and second component gives the total transportation cost.

(c) Let $z = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ and then compute $xRQz$, which is a number giving the total cost of materials and transportation for all the houses being built. [*Ans.* \$14,304.]

13. A college survey at an all-male school shows that dates of students are distributed as follows: a freshman dates one blonde and one brunette during the year; each sophomore dates one blonde, three brunettes, and one redhead; each junior dates three blondes, two brunettes, and two redheads; each senior dates three redheads. It is further known that each blonde brings three dresses with her, two skirts, two blouses, and one sweater; each brunette brings five dresses, four skirts, one blouse, and three sweaters; each redhead brings one dress, four skirts, and four sweaters. If each dress costs \$50, each skirt \$15, each blouse \$10, and each sweater \$5; and if there are 500 freshmen, 400 sophomores, 300 juniors, and 200 seniors,

(a) What is the total number of blondes, brunettes, and redheads dated?

(b) What is the total number of each type of clothing item in the dates' wardrobes?

(c) What is the cost of the wardrobe of a blonde? a brunette? a redhead?

(d) What is the total cost of all the wardrobes of all the dates? Calculate two ways. [*Ans.* \$1,347,500.]

5. THE SOLUTION OF LINEAR EQUATIONS

There are many occasions when the simultaneous solutions of linear equations is important. In this section we shall develop methods for finding out whether a set of linear equations has solutions, and for finding all such solutions.

Example 1. Consider the following example of three linear equations in three unknowns:

$$x_1 + 4x_2 + 3x_3 = 1 \tag{1}$$
$$2x_1 + 5x_2 + 4x_3 = 4 \tag{2}$$
$$x_1 - 3x_2 - 2x_3 = 5. \tag{3}$$

Before we discuss the solution of these equations we note that they can be written as a single equation in matrix form as follows:

$$\begin{pmatrix} 1 & 4 & 3 \\ 2 & 5 & 4 \\ 1 & -3 & -2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 4 \\ 5 \end{pmatrix}.$$

One of the uses of vector and matrix notation is in writing a large number of linear equations in a single simple matrix equation such as the one above.

The method of solving the linear equations above is the following. First we use equation (1) to eliminate the variable x_1 from equations (2) and (3); i.e., we subtract 2 times (1) from (2) and then subtract (1) from (3), giving

$$x_1 + 4x_2 + 3x_3 = 1 \tag{1'}$$
$$-3x_2 - 2x_3 = 2 \tag{2'}$$
$$-7x_2 - 5x_3 = 4. \tag{3'}$$

Next we divide equation (2') through by the coefficient of x_2, namely, -3, obtaining $x_2 + \frac{2}{3}x_3 = -\frac{2}{3}$. We use this equation to eliminate x_2 from each of the other two equations. In order to do this we subtract

4 times this equation from (1′) and add 7 times this equation to (3′), obtaining

$$(1'') \qquad x_1 + 0 + \tfrac{1}{3}x_3 = \tfrac{11}{3}$$
$$(2'') \qquad x_2 + \tfrac{2}{3}x_3 = -\tfrac{2}{3}$$
$$(3'') \qquad -\tfrac{1}{3}x_3 = -\tfrac{2}{3}.$$

The last step is to divide through (3″) by $-\frac{1}{3}$, which is the coefficient of x_3, obtaining the equation $x_3 = 2$; we use this equation to eliminate x_3 from the first two equations as follows:

$$(1''') \qquad x_1 + 0 + 0 = 3$$
$$(2''') \qquad x_2 + 0 = -2$$
$$(3''') \qquad x_3 = 2.$$

The solution can now be read from these equations as $x_1 = 3$, $x_2 = -2$, and $x_3 = 2$. The reader should substitute these values into the original equations (1), (2), and (3) above to see that the solution has actually been obtained.

In the example just discussed we saw that there was only one solution to the set of three simultaneous equations in three variables. Example 2 will be one in which there is *more* than one solution, and Example 3 will be one in which there are *no* solutions to a set of three simultaneous equations in three variables.

Example 2. Consider the following linear equations.

$$(4) \qquad x_1 - 2x_2 - 3x_3 = 2$$
$$(5) \qquad x_1 - 4x_2 - 13x_3 = 14$$
$$(6) \qquad -3x_1 + 5x_2 + 4x_3 = 0.$$

Let us proceed as before and use equation (1) to eliminate the variable x_1 from the other two equations. We have

$$(4') \qquad x_1 - 2x_2 - 3x_3 = 2$$
$$(5') \qquad -2x_2 - 10x_3 = 12$$
$$(6') \qquad -x_2 - 5x_3 = 6.$$

Proceeding as before, we divide equation (5′) by -2, obtaining the equation $x_2 + 5x_3 = -6$. We use this equation to eliminate the variable x_2 from each of the other equations—namely, we add twice this equation to (4′) and then add the equation to (6′).

$$x_1 + 0 + 7x_3 = -10 \quad (4'')$$
$$x_2 + 5x_3 = -6 \quad (5'')$$
$$0 = 0. \quad (6'')$$

Observe that we have eliminated the last equation completely! We also see that the variable x_3 can be chosen completely arbitrarily in these equations. To emphasize this, we move the terms involving x_3 to the right-hand side, giving

$$x_1 = -10 - 7x_3 \quad (4''')$$
$$x_2 = -6 - 5x_3. \quad (5''')$$

The reader should check, by substituting these values of x_1 and x_2 into equations (4), (5), and (6), that they are solutions regardless of the value of x_3. Let us also substitute particular values for x_3 to obtain numerical solutions. Thus, if we let $x_3 = 1, 0, -2$, respectively, and compute the resulting numbers, using (4‴) and (5‴), we obtain the following numerical solutions:

$$x_1 = -17, \quad x_2 = -11, \quad x_3 = 1$$
$$x_1 = -10, \quad x_2 = -6, \quad x_3 = 0$$
$$x_1 = 4, \quad x_2 = 4, \quad x_3 = -2.$$

The reader should also substitute these numbers into (4), (5), and (6) to show that they are solutions. To summarize, our second example has an infinite number of solutions, one for each numerical value of x_3 which is substituted into equations (4‴) and (5‴).

Example 3. Suppose that we modify equation (6) by changing the number on the right-hand side to 2. Then we have

$$x_1 - 2x_2 - 3x_3 = 2 \quad (7)$$
$$x_1 - 4x_2 - 13x_3 = 14 \quad (8)$$
$$-3x_1 + 5x_2 + 4x_3 = 2. \quad (9)$$

If we carry out the same procedure as before and use (7) to eliminate x_1 from (8) and (9), we obtain

$$x_1 - 2x_2 - 3x_3 = 2 \quad (7')$$
$$-2x_2 - 10x_3 = 12 \quad (8')$$
$$-x_2 - 5x_3 = 8. \quad (9')$$

We divide (8′) by -2, the coefficient of x_2, obtaining, as before,

$x_2 + 5x_3 = -6$. Using this equation to eliminate x_2 from the other two equations, we have

$$x_1 + 0 + 7x_3 = -10 \quad (7'')$$
$$x_2 + 5x_3 = -6 \quad (8'')$$
$$0 = 2. \quad (9'')$$

Observe that the last equation is *false*. Because our elimination procedure has led to a false result we conclude that the equations (7), (8), and (9) have *no* solution. The student should always keep in mind that this possibility exists when considering simultaneous equations.

In the examples above the equations we considered had the same number of variables as equations. The next example has more variables than equations and the last has more equations than variables.

Example 4. Consider the following two equations in three variables:

$$-4x_1 + 3x_2 + 2x_3 = -2 \quad (10)$$
$$5x_1 - 4x_2 + x_3 = 3. \quad (11)$$

Using the elimination method outlined above, we divide (10) by -4, and then subtract 5 times the result from (11), obtaining

$$x_1 - \tfrac{3}{4}x_2 - \tfrac{1}{2}x_3 = \tfrac{1}{2} \quad (10')$$
$$-\tfrac{1}{4}x_2 + \tfrac{7}{2}x_3 = \tfrac{1}{2}. \quad (11')$$

Multiplying (11′) by -4 and using it to eliminate x_2 from (10′), we have

$$x_1 + 0 - 11x_3 = -1 \quad (10'')$$
$$x_2 - 14x_3 = -2. \quad (11'')$$

We can now let x_3 take on any value whatsoever and solve these equations for x_1 and x_2. We emphasize this fact by rewriting them as in Example 2 as

$$x_1 = 11x_3 - 1 \quad (10''')$$
$$x_2 = 14x_3 - 2. \quad (11''')$$

The reader should check that these are solutions and also, by choosing specific values for x_3, find numerical solutions to these equations.

Example 5. Let us consider the other possibility suggested by Example 4, namely, the case in which we have more equations than variables. Consider the following equations:

$$-4x_1 + 3x_2 = 2 \tag{12}$$
$$5x_1 - 4x_2 = 0 \tag{13}$$
$$2x_1 - x_2 = a, \tag{14}$$

where a is an arbitrary number. Using equation (12) to eliminate x_1 from the other two we obtain

$$x_1 - \tfrac{3}{4}x_2 = -\tfrac{1}{2} \tag{12'}$$
$$-\tfrac{1}{4}x_2 = \tfrac{5}{2} \tag{13'}$$
$$\tfrac{1}{2}x_2 = a + 1. \tag{14'}$$

Next we use (13') to eliminate x_2 from the other equations, obtaining

$$x_1 + 0 = -8 \tag{12''}$$
$$x_2 = -10 \tag{13''}$$
$$0 = a + 6. \tag{14''}$$

These equations remind us of the situation in Example 3, since we will be led to a false result unless $a = -6$. We see that equations (12), (13), and (14) have the solution $x_1 = -8$ and $x_2 = -10$ only if $a = -6$. If $a \neq -6$, then there is *no* solution to these equations.

The examples above illustrate all the possibilities that can occur in the general case. There may be no solutions, exactly one solution, or an infinite number of solutions to a set of simultaneous equations.

The procedure that we have illustrated above is one that turns any set of linear equations into an equivalent set of equations from which the existence of solutions and the solutions can be easily read. A student who learned other ways of solving linear equations may wonder why we use the above procedure—one which is not always the quickest way of solving equations. The answer is that we use it because it always works, that is, it is a *canonical* procedure to apply to *any* set of linear equations. The faster methods usually work only for equations that have solutions, and even then they may not find all solutions. The value of a standard infallible method, especially for machine computation, should not be underestimated.

EXERCISES

1. Find all the solutions of the following simultaneous equations.

(a)
$$\begin{aligned} 4x_1 + 5x_3 &= 6. \\ x_2 - 6x_3 &= -2. \\ 3x_1 + 4x_3 &= 3. \end{aligned}$$
[*Ans.* $x_1 = 9, x_2 = -38, x_3 = -6.$]

(b)
$$\begin{aligned} 3x_1 - x_2 - 2x_3 &= 2. \\ 2x_2 - x_3 &= -1. \\ 3x_1 - 5x_2 &= 3. \end{aligned}$$
[*Ans.* No solution.]

(c)
$$\begin{aligned} -x_1 + 2x_2 + 3x_3 &= 0. \\ x_1 - 4x_2 - 13x_3 &= 0. \\ -3x_1 + 5x_2 + 4x_3 &= 0. \end{aligned}$$
[*Ans.* $x_1 = -7x_3, x_2 = -5x_3.$]

2. Find all the solutions of the following simultaneous equations.

(a)
$$\begin{aligned} x_1 + x_2 + x_3 &= 0. \\ 2x_1 + 4x_2 + 3x_3 &= 0. \\ 4x_2 + 4x_3 &= 0. \end{aligned}$$

(b)
$$\begin{aligned} x_1 + x_2 + x_3 &= -2. \\ 2x_1 + 4x_2 + 3x_3 &= 3. \\ 4x_2 + 2x_3 &= 2. \end{aligned}$$

(c)
$$\begin{aligned} 4x_1 + 4x_3 &= 8. \\ x_2 - 6x_3 &= -3. \\ 3x_1 + x_2 - 3x_3 &= 3. \end{aligned}$$

3. Find numbers x_1, x_2, and x_3 that solve the equations given in Exercise 1(c) and that also satisfy the nonlinear equation

$$x_1[2x_2 - 5x_3] = 1.$$

4. Find all solutions of the following equations:

(a)
$$\begin{aligned} 5x_1 - 3x_2 &= -7. \\ -2x_1 + 9x_2 &= 4. \\ 2x_1 + 4x_2 &= -2. \end{aligned}$$
[*Ans.* $x_1 = -\frac{17}{13}; x_2 = \frac{2}{13}.$]

(b)
$$\begin{aligned} x_1 + 2x_2 &= 1. \\ -3x_1 + 2x_2 &= -2. \\ 2x_1 + 3x_2 &= 1. \end{aligned}$$
[*Ans.* No solution.]

(c)
$$\begin{aligned} 5x_1 - 3x_2 - 7x_3 + x_4 &= 10. \\ -x_1 + 2x_2 + 6x_3 - 3x_4 &= -3. \\ x_1 + x_2 + 4x_3 - 5x_4 &= 0. \end{aligned}$$

5. Show that the equations

$$\begin{aligned} -4x_1 + 3x_2 + ax_3 &= c \\ 5x_1 - 4x_2 + bx_3 &= d \end{aligned}$$

always have a solution for all values of a, b, c, and d.

6. Find conditions on a, b, and c in order that the equations

$$\begin{aligned} -4x_1 + 3x_2 &= a \\ 5x_1 - 4x_2 &= b \\ -3x_1 + 2x_2 &= c \end{aligned}$$

have a solution. [*Ans.* $2a + b = c$.]

7. (a) Let $x = (x_1,x_2)$ and let A be the matrix

$$A = \begin{pmatrix} 3 & -4 \\ 2 & -6 \end{pmatrix}.$$

Find all solutions of the equation $xA = x$. [*Ans.* $x = (0, 0)$.]

(b) Let $x = (x_1,x_2)$ and let A be the matrix

$$A = \begin{pmatrix} 3 & 6 \\ -2 & -5 \end{pmatrix}.$$

Find all solutions of the equation $xA = x$.

[*Ans.* $x = (k,k)$ for any number k.]

8. Let $x = (x_1,x_2)$ and let P be the matrix

$$P = \begin{pmatrix} \frac{1}{3} & \frac{2}{3} \\ \frac{4}{5} & \frac{1}{5} \end{pmatrix}.$$

(a) Find all solutions of the equation $xP = x$.

(b) Choose the solution for which $x_1 + x_2 = 1$.

9. If $x = (x_1,x_2,x_3)$ and A is the matrix

$$A = \begin{pmatrix} 1 & -2 & 0 \\ 0 & 5 & 4 \\ 0 & -6 & -4 \end{pmatrix},$$

find all solutions of the equation $xA = x$.

[*Ans.* $x = (-k/2,5k/4,k)$ for any number k.]

10. If $x = (x_1,x_2,x_3)$ and P is the matrix

$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \\ \frac{1}{5} & 0 & \frac{4}{5} \end{pmatrix},$$

find all solutions of the equation $xP = x$. Select the unique solution for which $x_1 + x_2 + x_3 = 1$.

11. Find all solutions of:

$$\begin{aligned} x_1 + 2x_2 + 3x_3 + 4x_4 &= 10. \\ 2x_1 - x_2 + x_3 - x_4 &= 1. \\ 3x_1 + x_2 + 4x_3 + 3x_4 &= 11. \\ -2x_1 + 6x_2 + 4x_3 + 10x_4 &= 18. \end{aligned}$$

[*Ans.* $x_1 = \frac{12}{5} - x_3 - 2x_4/5$; $x_2 = \frac{19}{5} - x_3 - 9x_4/5$.]

12. We consider buying three kinds of food. Food I has one unit of vitamin A, three units of vitamin B, and four units of vitamin C. Food II has two, three and five units, respectively. Food III has three units each of vitamin A and vitamin C, none of vitamin B. We need to have 11 units of vitamin A, 9 of vitamin B, and 20 of vitamin C.

(a) Find all possible amounts of the three foods that will provide precisely these amounts of the vitamins.

(b) If Food I costs 60 cents and the others cost 10 cents each per unit, is there a solution costing exactly \$1? [*Ans.* (b) Yes; 1,2,2.]

6. THE INVERSE OF A SQUARE MATRIX

If A is a square matrix and B is another square matrix of the same size having the property that $BA = I$ (where I is the identity matrix), then we say that B is the *inverse* of A. When it exists we shall denote the inverse of A by the symbol A^{-1}. To give a numerical example, let A and A^{-1} be the following:

$$A = \begin{pmatrix} 4 & 0 & 5 \\ 0 & 1 & -6 \\ 3 & 0 & 4 \end{pmatrix},$$

$$A^{-1} = \begin{pmatrix} 4 & 0 & -5 \\ -18 & 1 & 24 \\ -3 & 0 & 4 \end{pmatrix}.$$

Then we have

$$A^{-1}A = \begin{pmatrix} 4 & 0 & -5 \\ -18 & 1 & 24 \\ -3 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 & 5 \\ 0 & 1 & -6 \\ 3 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I.$$

If we multiply these matrices in the other order we also get the identity matrix; thus

$$AA^{-1} = \begin{pmatrix} 4 & 0 & 5 \\ 0 & 1 & -6 \\ 3 & 0 & 4 \end{pmatrix} \cdot \begin{pmatrix} 4 & 0 & -5 \\ -18 & 1 & 24 \\ -3 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I.$$

In general it can be shown that if A is a square matrix with inverse A^{-1}, then the inverse satisfies the equation

$$A^{-1}A = AA^{-1} = I.$$

It is easy to see that a square matrix can have only one inverse. Suppose that in addition to A^{-1} we also have a B such that

$$BA = I.$$

Then we see that

$$B = BI = B(AA^{-1}) = (BA)A^{-1} = IA^{-1} = A^{-1}.$$

Finding the inverse of a matrix is analogous to finding the reciprocal of an ordinary number, but the analogy is not complete. Every non-zero number has a reciprocal, but there are matrices, not the zero matrix, which have no inverse. For example, if

$$A = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \quad \text{and} \quad B = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$$

then

$$AB = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} = 0.$$

From this it follows that neither A nor B can have an inverse. To show that A does not have an inverse, let us assume that A had an inverse A^{-1}. Then

$$B = (A^{-1}A)B = A^{-1}(AB) = A^{-1}0 = 0$$

contradicting the fact that $B \neq 0$. The proof that B cannot have an inverse is similar.

Let us try to calculate the inverse of a 2×2 matrix A; that is, let us find conditions on a matrix B that $BA = I$. Suppose

$$I = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} \cdot \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

$$= \begin{pmatrix} b_{11}a_{11} + b_{12}a_{21} & b_{11}a_{12} + b_{12}a_{22} \\ b_{21}a_{11} + b_{22}a_{21} & b_{21}a_{12} + b_{22}a_{22} \end{pmatrix}.$$

For these equalities to hold, the following equations must be satisfied

(1) $b_{11}a_{11} + b_{12}a_{21} = 1$

(2) $b_{11}a_{12} + b_{12}a_{22} = 0$

(3) $b_{21}a_{11} + b_{22}a_{21} = 0$

(4) $b_{21}a_{12} + b_{22}a_{22} = 1.$

In these equations the b's are to be regarded as the unknowns and the a's as constants. Observe that the unknowns b_{11} and b_{12} appear only in equations (1) and (2), while the unknowns b_{21} and b_{22} occur only in equations (3) and (4). In order to eliminate the variable b_{12} from equations (1) and (2) we multiply equation (1) by a_{22} and equation (2) by a_{21}, giving

$$\begin{aligned} b_{11}a_{11}a_{22} + b_{12}a_{21}a_{22} &= a_{22} \\ b_{11}a_{12}a_{21} + b_{12}a_{22}a_{21} &= 0. \end{aligned}$$

Subtracting the second of these equations from the first, we obtain

$$b_{11}(a_{11}a_{22} - a_{12}a_{21}) = a_{22}.$$

Now, providing the coefficient of b_{11} in this equation is not zero, we can divide through by it and obtain the solution for b_{11} as follows:

$$b_{11} = \frac{a_{22}}{a_{11}a_{22} - a_{21}a_{12}}.$$

Proceeding in this way for the other variables we can solve for the other variables, obtaining

$$b_{12} = \frac{-a_{12}}{a_{11}a_{22} - a_{21}a_{12}},$$

$$b_{21} = \frac{-a_{21}}{a_{11}a_{22} - a_{21}a_{12}},$$

$$b_{22} = \frac{a_{11}}{a_{11}a_{22} - a_{21}a_{12}}.$$

Notice that the denominator of each of these expressions is the same. This quantity is so important that it is given a special name, namely, the *determinant* of A, and it is defined as follows:

$$\det A = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{21}a_{12}.$$

Notice that we use vertical bars to denote a determinant, while we use vertical parentheses to denote a matrix.

Numerical examples of determinants are

$$\begin{vmatrix} 1 & 2 \\ 3 & 4 \end{vmatrix} = 4 - 6 = -2,$$

$$\begin{vmatrix} 1 & -1 \\ -1 & 1 \end{vmatrix} = 1 - 1 = 0.$$

We have defined determinants only for 2×2 matrices. They can also be defined for higher-order square matrices, but we shall not carry it out here.

Observe that we were able to solve for the inverse of a 2×2 matrix only if its determinant was not zero. Using the formulas derived previously the inverse of the matrix,

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$$

is

$$A^{-1} = \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix}.$$

To check this we observe that

$$\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} -2 & 1 \\ \frac{3}{2} & -\frac{1}{2} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

On the other hand, the matrix $\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ has determinant zero, so that the equations considered previously *cannot* be solved, and this matrix has no inverse. The same is true in general, and the following theorem can be stated.

Theorem. A square matrix has an inverse if and only if its determinant is nonzero.

The theorem holds for higher-order square matrices as well, but we will not prove it here.

EXERCISES

1. Find the determinants of the following matrices.

(a) $\begin{pmatrix} 1 & 6 \\ 0 & 1 \end{pmatrix}$. [*Ans.* 1.]

(b) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

(c) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$. [*Ans.* -1.]

(d) $\begin{pmatrix} 0 & 1 \\ 1 & 6 \end{pmatrix}$.

(e) $\begin{pmatrix} 2 & -3 \\ 4 & -6 \end{pmatrix}$. [*Ans.* 0.]

(f) $\begin{pmatrix} 2 & 2 \\ -4 & -4 \end{pmatrix}$.

(g) $\begin{pmatrix} -1 & 3 \\ 2 & 7 \end{pmatrix}$. [*Ans.* -13.]

2. Find when possible the inverses of the matrices given in Exercise 1. Check your work.

3. Let A be a square matrix that has an inverse. Show that the equations $Ax = b$ have $x = A^{-1}b$ as a solution.

4. Use the result of Exercise 3 to solve the equations

$$x_1 + 2x_2 = 1$$
$$3x_1 + 4x_2 = 2.$$

(*Hint:* The inverse of the matrix was computed in the text above.)

[*Ans.* $x_1 = 0$, $x_2 = \frac{1}{2}$.]

5. Let A be one of the matrices in Exercise 1, and let

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \quad \text{and} \quad b = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Using (when they exist) the inverses computed in Exercise 2, solve the equations $Ax = b$. (See Exercise 3.)

6. If $ad - cb \neq 0$, find the inverse of the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$.

7. Let A be the matrix of Exercise 6. Let

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \text{ and let } f = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}.$$

Then use the inverse found in Exercise 6 to solve the linear equations $Ax = f$.

8. If A is a square matrix that has an inverse A^{-1}, show that the inverse of A^2 is the matrix $[A^{-1}]^2 = A^{-2}$. Show that the inverse of A^n is the matrix $[A^{-1}]^n = A^{-n}$.

9. Solve the linear equations

$$\begin{aligned} 4x + 5z &= 7 \\ y - 6z &= 2 \\ 3x + 4z &= -1 \end{aligned}$$

by writing them in the form $Ax = b$ and using the inverse of A given in the beginning of this section. (See Exercise 3.)

[*Ans.* $x = 33$, $y = -148$, $z = -25$.]

10. (a) Prove the following identity for determinants:

$$\left|\begin{pmatrix} a & b \\ c & d \end{pmatrix} \cdot \begin{pmatrix} e & f \\ g & h \end{pmatrix}\right| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} \cdot \begin{vmatrix} e & f \\ g & h \end{vmatrix}.$$

(b) Show by an example that it is *not* true in general that the determinant of the sum of two matrices is equal to the sum of their determinants.

11. Use the result of Exercise 10(a) to show that the determinant of the product of two matrices can be zero only if one of the matrices itself has a zero determinant.

12. If A and B are $n \times n$ matrices each of which has an inverse:

(a) Prove, using Exercise 11, that AB has an inverse.

(b) Express the inverse of AB in terms of A^{-1} and B^{-1}.

7. APPLICATIONS OF MATRIX THEORY TO MARKOV CHAINS

For simplicity we shall confine our discussion to three-state Markov chains, but a similar procedure will work for any other Markov chain.

In Section 13 of Chapter IV, we noted that to each Markov chain there was a matrix of transition probabilities. For example, if there are three states, a_1, a_2, and a_3, then

$$P = \begin{array}{c} \\ a_1 \\ a_2 \\ a_3 \end{array} \begin{array}{c} \begin{array}{ccc} a_1 & a_2 & a_3 \end{array} \\ \begin{pmatrix} p_{11} & p_{12} & p_{13} \\ p_{21} & p_{22} & p_{23} \\ p_{31} & p_{32} & p_{33} \end{pmatrix} \end{array}$$

is the transition matrix for the chain. Recall that the *row sums* of P are all equal to 1. Such a matrix is called a stochastic matrix.

DEFINITION. A *stochastic matrix* is a square matrix with nonnegative entries such that the sum of the entries in each row is 1.

In order to obtain a Markov chain we must specify how the process starts. Suppose that the initial state is chosen by a chance device that selects state a_j with probability $p_j^{(0)}$. We can represent these initial probabilities by means of the vector $p^{(0)} = (p_1^{(0)}, p_2^{(0)}, p_3^{(0)})$. As in Exercise 10 of Section 4, we can construct a tree measure for as many steps of the process as we wish to consider. Let $p_j^{(n)}$ be the probability that the process will be in state a_j after n steps. Let the vector of these probabilities be $p^{(n)} = (p_1^{(n)}, p_2^{(n)}, p_3^{(n)})$.

DEFINITION. A row vector p is called a *probability vector* if it has nonnegative components whose sum is 1.

Obviously the vectors $p^{(0)}$ and $p^{(n)}$ are probability vectors. Also each row of a stochastic matrix is a probability vector.

By means of the tree measure it can be shown that these probabilities satisfy the following equations:

$$\begin{aligned} p_1^{(n)} &= p_1^{(n-1)}p_{11} + p_2^{(n-1)}p_{21} + p_3^{(n-1)}p_{31} \\ p_2^{(n)} &= p_1^{(n-1)}p_{12} + p_2^{(n-1)}p_{22} + p_3^{(n-1)}p_{32} \\ p_3^{(n)} &= p_1^{(n-1)}p_{13} + p_2^{(n-1)}p_{23} + p_3^{(n-1)}p_{33}. \end{aligned}$$

It is not hard to give intuitive meanings to these equations. The first one, for example, expresses the fact that the probability of being in state a_1 after n steps is the sum of the probabilities of being at each of the three possible states after $n - 1$ steps and then moving to state a_1 on the nth step. The interpretation of the other equations is similar.

If we recall the definition of the product of a vector times a matrix we can write the above equations as

$$p^{(n)} = p^{(n-1)}P.$$

If we substitute small values of n we get the equations: $p^{(1)} = p^{(0)}P$; $p^{(2)} = p^{(1)}P = p^{(0)}P^2$; $p^{(3)} = p^{(2)}P = p^{(0)}P^3$; etc. In general, it can be seen that

$$p^{(n)} = p^{(0)}P^n.$$

Thus we see that, if we multiply the vector $p^{(0)}$ of initial probabilities by the nth power of the transition matrix P, we obtain the vector $p^{(n)}$, whose components give the probabilities of being in each of the states after n steps.

In particular, let us choose $p^{(0)} = (1,0,0)$ which is equivalent to letting the process start in state a_1. From the equation above we see that then $p^{(n)}$ is the first row of the matrix P^n. Thus the elements of the first row of the matrix P^n give us the probabilities that after n steps the process will be in a given one of the states, under the assumption that it started in state a_1. In the same way, if we choose $p^{(0)} = (0,1,0)$, we see that the second row of P^n gives the probabilities that the process will be in one of the various states after n steps, given

that it started in state a_2. Similarly the third row gives these probabilities, assuming that the process started in state a_3.

In Section 13 of Chapter IV, we considered special Markov chains that started in given fixed states. There we arrived at a matrix $P^{(n)}$ whose ith row gave the probabilities of the process ending in the various states, given that it started at state a_i. By comparing the work that we did there with what we have just done, we see that the matrix $P^{(n)}$ is merely the nth power of P, that is, $P^{(n)} = P^n$. (Compare Exercise 10 of Section 4.) Matrix multiplication thus gives a convenient way of computing the desired probabilities.

The equation $p^{(n)} = p^{(n-1)}P$ admits of another interesting interpretation. The vector $p^{(n)}$ is obtained from the vector $p^{(n-1)}$ by a transformation consisting of multiplying it by the matrix P. The vector $p^{(n-1)}$ is obtained from $p^{(n-2)}$ by a similar transformation, etc. Let us indicate the transformation by an arrow; thus we have

$$p^{(0)} \to p^{(1)} \to p^{(2)} \to \ldots \to p^{(n-1)} \to p^{(n)} \to \ldots .$$

In Section 9 we shall show that this transformation is what is called a *linear transformation* of vectors. We shall say that the transformation *sends* the vector $p^{(0)}$ onto the vector $p^{(1)}$, and sends $p^{(1)}$ onto $p^{(2)}$, etc.

Sometimes it happens that there is a probability vector t which is sent by the transformation P onto itself, that is, $t = tP$. If we interpret t as a point in Euclidean space, then we say that t is a *fixed point* of the transformation P.

DEFINITION. The probability vector t is a *fixed point* of the transformation P, if $t = tP$.

Example. Consider the stochastic matrix

$$P = \begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} .667 & .333 \\ .500 & .500 \end{pmatrix}.$$

If $t = (.6,.4)$, then we see that

$$tP = (.6,.4)\begin{pmatrix} \frac{2}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix} = (.6,.4) = t,$$

so that t is the fixed point of the transformation t.

If we had happened to choose the vector t as our initial probability vector $p^{(0)}$, we would have had $p^{(n)} = p^{(0)}P^n = tP^n = t = p^{(0)}$. In this case the probability of being at any particular state is the same at all steps of the process. Such a process we shall call a *stationary Markov process.*

As seen above, in the study of Markov chains we are interested in the powers of the matrix P. To see what happens to these powers, let us further consider the example.

Example (continued). Suppose that we compute powers of the matrix P in the example above. We have

$$P^2 = \begin{pmatrix} .611 & .389 \\ .583 & .417 \end{pmatrix}, \quad P^3 = \begin{pmatrix} .602 & .398 \\ .597 & .403 \end{pmatrix}, \quad \text{etc.}$$

It looks as if the matrix P^n is approaching the matrix

$$T = \begin{pmatrix} .6 & .4 \\ .6 & .4 \end{pmatrix}$$

and, in fact, it can be shown that this is the case. (When we say that P^n approaches T we mean that each entry in the matrix P^n gets close to the corresponding entry in T.) Note that each row of T is the fixed point t of the matrix P.

We cannot prove here that P^n approaches T, so we shall content ourselves with stating a useful result. For that we need the definition of a regular matrix.

DEFINITION. A stochastic matrix is said to be *regular* if some power of the matrix has only positive components.

Thus the matrix in the example is regular, since every entry in it is positive, so that the first power of the matrix has all positive entries. Other examples occur in the exercises.

Theorem. If P is a regular stochastic matrix, then

(i) The powers P^n approach a matrix T.
(ii) Each row of T is the same probability vector t.
(iii) The components of t are positive.

We omit the proof of this theorem; however we can prove the next theorem.

Theorem. If P is a regular stochastic matrix, and T and t are given by the previous theorem, then

(a) If p is any probability vector, pP^n approaches t.

(b) The vector t is the unique fixed point probability vector of P.

Proof. First let us consider the vector pT. The first column of T has a t_1 in each row. Hence in the first component of pT each component of p is multiplied by t_1, and therefore we have t_1 times the sum of the components of p, which is t_1. Doing the same for the other components, we note that pT is simply t. But pP^n approaches pT; hence it approaches t. Thus if any probability vector is transformed repeatedly by P, it approaches the fixed point t. This proves part (a).

Since the powers of P approach T, $P^{n+1} = P^nP$ approaches T, but it also approaches TP; hence $TP = T$. Any one row of this matrix equation states that $tP = t$; hence t is a fixed point (and by the previous theorem a probability vector). We must still show that it is unique. Let u be any probability vector fixed point of P. By part (a) we know that uP^n approaches t. But since u is a fixed point, $uP^n = u$. Hence u remains fixed but "approaches" t. This is possible only if $u = t$. Hence t is the only probability vector fixed point. This completes the proof of part (b).

The following is an important consequence of this theorem. If we take as p the vector $p^{(0)}$ of initial probabilities, then the vector $pP^n = p^{(n)}$ gives the probabilities after n steps, and this vector approaches t. Therefore, no matter what the initial probabilities are, if P is regular, then after a large number of steps the probability that the process is in state a_j will be very nearly t_j.

Example (continued). Let us take $p^{(0)} = (.1,.9)$ and see how the successive transformations of it change. Using P as in the example above, we have that $p^{(1)} = (.5167,.4833)$, $p^{(2)} = (.5861,.4139)$, and $p^{(3)} = (.5977,.4023)$. Recalling that $t = (.6,.4)$, we see that these vectors do approach t. They are plotted in Figure 5.

As a final example let us derive the formulas for the fixed point of a 2×2 stochastic matrix with positive components. Such a matrix is of the form

$$S = \begin{pmatrix} 1 - a & a \\ b & 1 - b \end{pmatrix}$$

where $0 < a < 1$ and $0 < b < 1$. Since S is regular, it has a unique probability vector fixed point $t = (t_1,t_2)$. Its components must satisfy the equations

$$t_1(1 - a) + t_2 b = t_1$$
$$t_1 a + t_2(1 - b) = t_2.$$

Each of these equations reduces to the single equation $t_1 a = t_2 b$. This single equation has an infinite number of solutions. However, since t is a probability vector, we must also have $t_1 + t_2 = 1$, and the new equation gives the point $[b/(a+b),a/(a+b)]$ as the unique fixed-point probability vector of S.

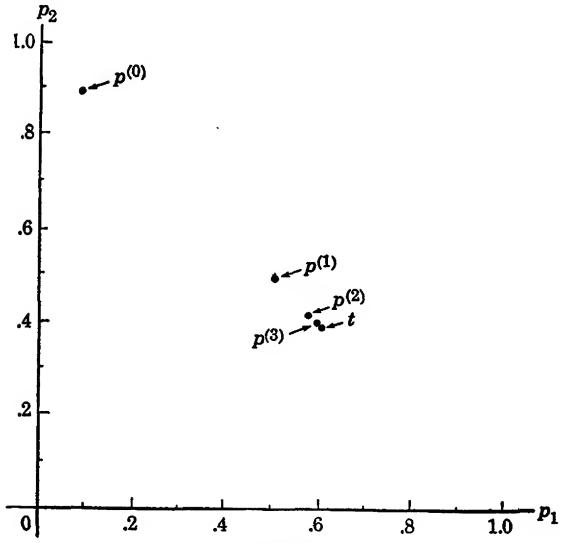


Figure 5

EXERCISES

1. Which of the following matrices are regular?

(a) $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$ (b) $\begin{pmatrix} 0 & 1 \\ \frac{1}{4} & \frac{3}{4} \end{pmatrix}$ [Regular]

(c) $\begin{pmatrix} 1 & 0 \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$ (d) $\begin{pmatrix} \frac{1}{5} & \frac{4}{5} \\ 1 & 0 \end{pmatrix}$ [Regular]

(e) $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ 0 & 1 \end{pmatrix}$ (f) $\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}$ [Not regular]

(g) $\begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}$ (h) $\begin{pmatrix} \frac{1}{3} & 0 & \frac{2}{3} \\ 0 & 1 & 0 \\ 0 & \frac{1}{5} & \frac{4}{5} \end{pmatrix}$. [Not regular]

2. Show that the 2×2 matrix

$$S = \begin{pmatrix} 1 - a & a \\ b & 1 - b \end{pmatrix}$$

is a regular stochastic matrix if and only if either

(i) $0 < a \leq 1$ and $0 < b < 1$; or
(ii) $0 < a < 1$ and $0 < b \leq 1$.

3. Find the fixed point for the matrix in Exercise 2 for each of the cases listed there. (*Hint:* Most of the cases were covered in the text above.)

4. Find the fixed point t for each of the following regular matrices.

(a) $\begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}$. [*Ans.* $t = (\frac{2}{3}, \frac{1}{3})$.]

(b) $\begin{pmatrix} .9 & .1 \\ .1 & .9 \end{pmatrix}$.

(c) $\begin{pmatrix} \frac{3}{4} & \frac{1}{4} & 0 \\ 0 & \frac{2}{3} & \frac{1}{3} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{pmatrix}$. [*Ans.* $t = (\frac{2}{7}, \frac{3}{7}, \frac{2}{7})$.]

5. Let $p^{(0)} = (\frac{1}{2}, \frac{1}{2})$ and compute $p^{(1)}$, $p^{(2)}$, and $p^{(3)}$ for each of the matrices in Exercises 4(a) and 4(b). Do they approach the fixed points of these matrices?

6. Give a probability theory interpretation to the condition of regularity.

7. Consider the two-state Markov chain with transition matrix

$$P = \begin{matrix} \\ a_1 \\ a_2 \end{matrix} \begin{matrix} \begin{matrix} a_1 & a_2 \end{matrix} \\ \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \end{matrix}.$$

What is the probability that after n steps the process is in state a_1 if it started in state a_2? Does this probability become independent of the initial position for large n? If not, the theorem of this section must not apply. Why? Does the matrix have a unique fixed point probability vector?

8. Prove that, if a regular 3×3 transition matrix has the property that its column sums are 1, its fixed point probability vector is $(\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$. State a similar result for $n \times n$ transition matrices having column sums equal to 1.

9. Compute the first five powers of the matrix

$$P = \begin{pmatrix} .8 & .2 \\ .2 & .8 \end{pmatrix}.$$

From these, guess the fixed point vector t. Check by computing what t is.

10. Show that all stochastic matrices of the form

$$\begin{pmatrix} 1-a & a \\ a & 1-a \end{pmatrix}$$

where $0 < a < 1$ have the same unique fixed point. [*Ans.* $t = (\frac{1}{2},\frac{1}{2})$.]

11. In Exercise 9 take $p = (.7,.3)$ and compute pP, pP^2, pP^3, pP^4, and pP^5. Compare your results with the fixed vector t.

12. Compute the fixed vector of the matrix

$$S = \begin{pmatrix} .1 & .9 \\ .6 & .4 \end{pmatrix}.$$

Let $p = (.5,.5)$ and compute pS, pS^2, and pS^3. Plot these vectors as is done in Figure 5.

[*Ans.* $pS = (.35,.65)$, $pS^2 = (.425,.575)$. $pS^3 = (.3875,.6125)$, $t = (.4,.6)$.]

13. Let S be the matrix

$$S = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & \frac{1}{2} \end{pmatrix}.$$

Compute the unique probability vector fixed point of S, and use your result to prove that S is not regular.

14. Show that the matrix

$$S = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 1 \end{pmatrix}$$

has more than one probability vector fixed point. Find the matrix that S^n approaches, and show that it is not a matrix all of whose rows are the same.

15. Show that the matrix

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ \frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix}$$

is a regular matrix.

8. EXAMPLES OF MARKOV CHAINS

In this section we shall apply the results of the last section to several different Markov chains. Some of these will be new Markov chains, and some will be taken from the examples in Chapter 4, Section 13.

Example 1. Suppose that the President of the United States tells person A his intention either to run or not to run in the next election. Then A relays the news to B, who in turn relays the message to C, etc., always to some new person. Assume that there is a probability $p > 0$ that any one person, when he gets the message, will reverse it before passing it on to the next person. What is the probability that the nth man to hear the message will be told that the President will run? We can consider this as a two-state Markov chain, with states indicated by "yes" and "no." The process is in state "yes" at time n if the nth person to receive the message was told that the President would run. It is in state "no" if he was told that the President would not run. The matrix P of transition probabilities is then

$$\begin{array}{c} \\ \text{yes} \\ \text{no} \end{array} \begin{array}{c} \begin{array}{cc} \text{yes} & \text{no} \end{array} \\ \begin{pmatrix} 1-p & p \\ p & 1-p \end{pmatrix} \end{array}.$$

Then the matrix P^n gives the probabilities that the nth man is given a certain answer, assuming that the President said "yes" (first row) or assuming that the President said "no" (second row). We know that these rows approach t. From the formulas of the last section, we find that $t = (\frac{1}{2},\frac{1}{2})$. Hence the probabilities for the nth man being told "yes" or "no" approach $\frac{1}{2}$ independently of the initial decision of the President. For a large number of people, we can expect that approximately one-half will be told that the President will run and the other half that he will not, independently of the actual decision of the President.

Suppose now that the probability a that a person will change the news from "yes" to "no" when transmitting it to the next person is different from the probability b that he will change it from "no" to "yes." Then the matrix of transition probabilities becomes

$$\begin{array}{c} \\ \text{yes} \\ \text{no} \end{array} \begin{array}{c} \begin{array}{cc} \text{yes} & \text{no} \end{array} \\ \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \end{array}.$$

In this case $t = [b/(a+b),a/(a+b)]$. Thus there is a probability of approximately $b/(a+b)$ that the nth person will be told that the President will run. Assuming that n is large, this probability is independent of the actual decision of the President. For n large we can expect, in this case, that a proportion approximately equal to

$b/(a + b)$ will have been told that the President will run, and a proportion $a/(a + b)$ will have been told that he will not run. The important thing to note is that, from the assumptions we have made, it follows that it is not the President but the people themselves who determine the probability that a person will be told "yes" or "no," and the proportion of people in the long run that are given one of these predictions.

Example 2. For this example, we continue the study of Example 2 in Chapter IV, Section 13. The first approximation treated in that example leads to a two-state Markov chain, and the results are similar to those obtained in Example 1 above. The second approximation led to a four-state Markov chain with transition probabilities given by the matrix

$$\begin{array}{c} \\ \text{RR} \\ \text{DR} \\ \text{RD} \\ \text{DD} \end{array} \begin{array}{c} \begin{array}{cccc} \text{RR} & \text{DR} & \text{RD} & \text{DD} \end{array} \\ \begin{pmatrix} 1-a & 0 & a & 0 \\ b & 0 & 1-b & 0 \\ 0 & 1-c & 0 & c \\ 0 & d & 0 & 1-d \end{pmatrix} \end{array}.$$

If a, b, c, and d are all different from 0 or 1, then the square of the matrix has no zeros, and hence the matrix is regular. The fixed probability vector is found in the usual way (see Exercise 12) and is

$$\left(\frac{bd}{bd + 2ad + ca}, \frac{ad}{bd + 2ad + ca}, \frac{ad}{bd + 2ad + ca}, \frac{ca}{bd + 2ad + ca}\right).$$

Note that the probability of being in state RD after a large number of steps is equal to the probability of being in state DR. This shows that in equilibrium a change from R to D must have the same probability as a change from D to R.

From the fixed vector we can find the probability that an election in the far future will result in a victory for the Republicans. This is found by adding the probability of being in state RR and DR, giving

$$\frac{bd + ad}{bd + 2ad + ca}.$$

Notice that, to find the probability of a Republican victory on the year preceding some year far in the future, we should add the probabilities of being in states RR and RD. That we get the same result

corresponds to the fact that predictions far in the future are essentially independent of the particular year being predicted. In other words, the process is acting as if it were a stationary process.

Example 3. Here we continue the study of Example 3 in Chapter IV, Section 13. In that example, for any particular n we could write the transition matrix and solve for the fixed vector. However, it turns out to be more instructive to try and guess the answer. It would seem that after a large number of exchanges the balls should become pretty well scrambled up, and that the probability that any particular n balls should be in urn 1 should be the same as if we simply took n balls at random from $2n$ balls, n black and n white, and put them in the urn. If this were done, the probability that we would put j black balls in the urn is

$$p_j = \frac{\binom{n}{j}\binom{n}{n-j}}{\binom{2n}{n}} = \frac{\binom{n}{j}^2}{\binom{2n}{n}}.$$

It can be checked that these probabilities do satisfy the necessary equations for a fixed vector for the matrix, for any n.

Example 4. The following example shows how the above ideas can be applied to a situation which does not seem to be a probabilistic situation but which can be so interpreted.

Suppose that, in a certain city, each year four per cent of the people in the city (proper) move to the suburbs, and one per cent of the people in the suburbs move to the city. Assuming that the total number of people in the city plus its suburbs remains constant, what is the ultimate distribution of people between the city and suburbs? The matrix S of Figure 6 shows the situation.

	Move to city	Move to suburbs
People in city	.96	.04
People in suburbs	.01	.99

Figure 6

Let $x^{(0)}$ be the vector $(x_1^{(0)}, x_2^{(0)})$, where $x_1^{(0)}$ is the proportion of people in the city when the study begins and $x_2^{(0)}$ is the proportion in the

suburbs at this time. Then denote by $x^{(n)} = (x_1^{(n)}, x_2^{(n)})$ the vector giving the corresponding proportions after n years. Then, by the same argument as given for the probabilities in the case of a Markov chain, it follows that

$$x^{(n)} = x^{(0)}P^n.$$

Also, since our matrix P could be interpreted as a regular transition matrix for a Markov chain, and since the vector $x^{(0)}$ is a probability vector, we can apply the theorem from the theory of Markov chains and obtain the fact that the vector $x^{(n)}$ approaches the vector $t = (t_1, t_2)$, which is the unique probability vector fixed point of P. This vector is $t = (.2, .8)$. Thus we can conclude that after a long time there will be a fraction approximately equal to 20 per cent of the people in the city proper and 80 per cent of the people in the suburbs, independent of the fraction initially in the city and in the suburbs. Notice that, after a long time has elapsed, a fraction $(.2) \cdot (.04) = .008$ of the people move in a given year from the city to the suburbs and a fraction $(.8) \cdot (.01) = .008$ of the people move in a given year from the suburbs to the city. That is, in the "equilibrium" position which is reached after a long time, there are just as many people moving from the city to the suburbs as there are from the suburbs to the city.

Example 5. Suppose that in Example 4 we are interested in the following kind of problem. Assume that there is a known number of Republicans and Democrats initially in the city and in the suburbs. Assume also that there is no changing of party affiliations and that the decision to move at any given time is independent of the party affiliations. What then is the probability that after a long time there will be a particular party division in the city? (We know from the previous example that eventually the situation becomes one of simply transferring a fixed number of people from the city to the suburbs and the same number from the suburbs to the city in each year.) We shall indicate how this problem can be solved for a very simple special case.

Assume that there are two people in the city and eight in the suburbs when the equilibrium position is reached, and that each year one moves from the city to the suburbs and one from the suburbs to the city. Assume further that there are four Republicans and six Democrats in all. We form a Markov chain by taking as a state the

number of people in the city who are Republicans at any one time. Thus the states can be represented by the numbers 0, 1, and 2. The transition probabilities can be calculated as follows. Assume, for example, that we are in state 1. Then the situation is represented in Figure 7.

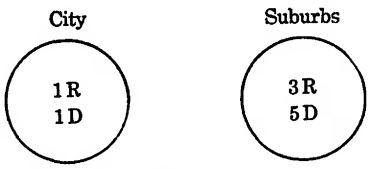


Figure 7

To go from state 1 to state 0 we must choose a Republican from the city and a Democrat from the suburbs. This happens with probability $\frac{1}{2} \cdot \frac{5}{8} = \frac{5}{16}$. To go from state 1 to state 1 we must choose either a Republican from each place or a Democrat from each place. The probability of this occurring is $\frac{1}{2} \cdot \frac{3}{8} + \frac{1}{2} \cdot \frac{5}{8} = \frac{1}{2}$. The other transition probabilities are calculated in a similar manner. We obtain the matrix of transition probabilities

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{5}{16} & \frac{1}{2} & \frac{3}{16} \\ 0 & \frac{3}{4} & \frac{1}{4} \end{pmatrix} \end{array}.$$

Our matrix is a regular matrix, since its square has all positive elements, and $t = (\frac{1}{3}, \frac{8}{15}, \frac{2}{15})$. (See Exercise 7.) Thus, after a long time the probability that the city will have no Republicans is approximately $\frac{1}{3}$, the probability for one Republican is $\frac{8}{15}$ and for two Republicans $\frac{2}{15}$. Again, these probabilities are independent of the initial number.

These limiting probabilities are the same that one obtains if one chooses two people at random from a group of ten, of whom four are Republicans and six are Democrats, and asks for the probability of obtaining zero, one, or two Republicans. (See Exercise 8.)

It is interesting to note that we ended up here with a chain like that in Example 3, which has been used by physicists as a crude model for diffusion of gases.

Example 6. Suppose that you are confronted with two slot machines each of which, when it pays off, pays off the same amount. You are told that machine A pays off each time with probability $\frac{1}{2}$, and that machine B pays off each time with probability $\frac{1}{4}$. If you are going to play one of the machines, you would naturally like to play A.

Unfortunately, you are not told which machine is which. We shall consider two systems which you might use and for each system find the fraction of times you can expect to win.

As a first system we assume that you decide to use only the result of the last play to make your decision whether to switch machines or to stay with the machine you played last time. For the first play we assume that you choose a machine at random. If you win on your first play, then the conditional probability that you have played machine A is $\frac{2}{3}$. (See Chapter IV, Section 7, Example 3.) Hence, if you win on the first play, you stay with the same machine next time. If you lose on the first play, the conditional probability that you have played machine A is $\frac{2}{5}$. Hence, you would switch machines in this case. The assumption that you use only the last result makes every play look like the first play, and hence you always use the decision, play the same machine next time if you won last time, and switch machines if you lost.

To find how you will fare under this system we form a Markov chain by taking as states the machines A and B. If you play machine A, the probability that you play A next time is $\frac{1}{2}$, i.e., the probability that you win. The other transition probabilities are found similarly, and we have the matrix of transition probabilities

$$\begin{array}{c} \\ \text{A} \\ \text{B} \end{array} \begin{array}{c} \begin{array}{cc} \text{A} & \text{B} \end{array} \\ \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{3}{4} & \frac{1}{4} \end{pmatrix}. \end{array}$$

The fixed vector of this matrix $(\frac{3}{5},\frac{2}{5})$ gives the probability after a number of plays that you will play each of the machines. Thus we see that, by this system, you can expect to play machine A about $\frac{3}{5}$ of the time. Hence you will win approximately $\frac{3}{5}\cdot\frac{1}{2} + \frac{2}{5}\cdot\frac{1}{4} = \frac{2}{5}$ of the time.

As a second system we assume that you use the outcome of the last two plays to determine whether to switch or not. We assume that after each decision you make two plays and on the basis of the outcome of these two plays make your next decision. In this case the conditional probability that you have machine A, given that you have made two plays resulting in win-win, is greater than $\frac{1}{2}$. This probability is also greater than $\frac{1}{2}$ for the case that plays resulted in win-lose or lose-win. (See Exercise 10.) Hence in each of these cases you stay with the same machine for your next two plays. In the case of lose-

lose, the probability that you have machine A is less than $\frac{1}{2}$, and you switch machines. Under this system of play the matrix of transition probabilities is

$$\begin{array}{c} \\ \text{A} \\ \text{B} \end{array} \begin{array}{c} \begin{array}{cc} \text{A} & \text{B} \end{array} \\ \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ \frac{9}{16} & \frac{7}{16} \end{pmatrix} \end{array},$$

where A means playing machine A twice. The fixed vector for this matrix is $(\frac{9}{13},\frac{4}{13})$. Thus you will by this system play machine A about $\frac{9}{13}$ of the time, and you will win a fraction of the time approximately equal to $\frac{9}{13}\cdot\frac{1}{2} + \frac{4}{13}\cdot\frac{1}{4} = \frac{11}{26} = .42$. Thus the more complicated system has improved your expectation by only .02. The best that any system could insure is an expectation of $\frac{1}{2}$.

EXERCISES

1. The land of Oz is blessed by many things, but not good weather. They *never* have two nice days in a row. If they have a nice day they are just as likely to have snow as rain the next day. If they have snow (or rain), they have an even chance of having the same the next day. If there is a change from snow or rain, only half of the time is this a change to a nice day. Set up a three-state Markov chain to describe this situation. Find the long-range probability for rain, for snow, and for a nice day. What fraction of the days does it rain in the land of Oz?

[*Ans.* The probabilities are: nice, $\frac{1}{5}$; rain, $\frac{2}{5}$; snow, $\frac{2}{5}$.]

2. In Example 2, assume that $a = \frac{1}{3}$, $b = c = \frac{1}{2}$, and $d = \frac{1}{4}$. Find the fixed vector. What proportion of future elections can be expected to be Republican victories under these assumptions?

3. In Example 2, assume that $a = 1$ and $b = 1$ and $c = d$, and that c is neither 0 nor 1. Show that the matrix is regular. Find the fixed vector. What limitation do our assumptions put on possible sequences of election outcomes? [*Ans.* $t = (\frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4})$.]

4. In Example 2, assume that $a = 0$ and b, c, and d are different from 0 or 1. Is the matrix regular? Show that (1,0,0,0) is a fixed vector. Interpret this vector in terms of long-range predictions.

5. In Example 2, assume that $a = 0$, $d = 0$, and b and c are different from 1 or 0. What can be said in this case about the nature of the political system after a long time?

[*Ans.* From some time on, the party in power remains in power.]

6. In Example 3, find the matrix of transition probabilities for the case $n = 2$, and show that the fixed vector agrees with the result guessed on intuitive grounds in the discussion of the example.

7. In Example 5, check that the fixed probability vector is $(\frac{1}{3},\frac{8}{15},\frac{2}{15})$.

8. Assume that from a group of ten people, four of whom are Republicans and six are Democrats, two are chosen at random. Find the probability that the two are both Republicans, that one is Republican and one is a Democrat, and that both are Democrats.

9. Assume that a certain salesman always goes from city A to city B, and always from city B to city C. However, from city C he goes with probability $\frac{1}{2}$ to city A and with probability $\frac{1}{2}$ to city B. Form a Markov chain to represent his travels. Is the matrix of transition probabilities regular? If so, find the fixed vector. [*Ans.* Yes. $t = (\frac{1}{5},\frac{2}{5},\frac{2}{5})$.]

10. Referring to the second system in Example 6, find the probability that the player has chosen machine A, given that the first two outcomes were win-win. Answer the same question for each of the other three possibilities from the first two outcomes.

[*Ans.* Win-win, $\frac{4}{5}$; win-lose, $\frac{4}{7}$; lose-win, $\frac{4}{7}$; lose-lose, $\frac{4}{13}$.]

11. In Example 6 assume that the player decides after every three plays which machine to play for the next three plays. He changes machines if and only if the conditional probability that he has machine A, given the last three outcomes, is less than one-half. Find the fraction of time that the player can expect to win using this system. Is this a better system than that based on only the last two times, given in Example 6? [*Ans.* 0.41; no.]

12. Show that the vector given in Example 2 is the fixed vector of the matrix of transition probabilities.

13. In Chapter IV, Section 13, Exercise 10, find the fixed point probability vector, and interpret it.

14. A professor tries not to be late for class too often. If he is late one day, he is 90 per cent sure to be on time next time. If he is on time, then the next day there is a 30 per cent chance of his being late. In the long run, how often is he late for class?

15. A professor has three pet questions, one of which occurs on every test he gives. The students know his habits well. He never uses the same question twice in a row. If he used question one last time, he tosses a coin, and uses question two if a head comes up. If he used question two, he tosses two coins and switches to question three if both come up heads. If he used question three, he tosses three coins and switches to question one if all three come up heads. In the long run, which question does he use most often, and how frequently is it used? [*Ans.* Question two, 40 per cent of the time.]

*9. LINEAR FUNCTIONS AND TRANSFORMATIONS

The primary use of vectors and matrices in science is the representation of several different quantities as a single one. For example, the demands on all the industries in the United States may be represented by a row vector x. We have seen examples where such a vector is multiplied by a column vector y, giving the number $x \cdot y$. The components of y could be the values of unit outputs of the various industries. Then $x \cdot y$ is the total monetary value of the demand on industries.

This illustration is typical of much that we meet in the sciences. It has two fundamental properties. If the demand increases by a given factor k, then $(kx) \cdot y = k(x \cdot y)$, and hence the value increases by the same factor. And if we have two demand vectors x and x', then $(x + x') \cdot y = (x \cdot y) + (x' \cdot y)$, and hence their values are also added.

Thus we see that y has the effect of assigning to each row vector x a number $f(x)$, and has the two very simple properties,

$$\text{(i)} \qquad f(kx) = kf(x)$$

$$\text{(ii)} \qquad f(x + x') = f(x) + f(x')$$

Such an assignment of a number to each row vector x we call a *linear function*. We have seen that each column vector with n components defines a linear function for row vectors with n components.

Linear functions represent the simplest type of dependence. Fortunately, very many problems can be represented at least approximately by linear functions. While it is not strictly true that manufacturing 100 tons of steel costs ten times as much as manufacturing 10 tons, this is at least a reasonable approximation. And the same holds for necessary raw materials, for labor needed, transportation costs, etc. Linear functions are so simple to handle that we try to use them whenever this is reasonable.

Not only is it true that every column vector represents a linear function, but every linear function of row vectors can be so represented. We will prove this for linear functions of three-component row vectors.

Let us suppose that f assigns a number $f(x)$ to each three-component vector x, and that it has the properties (i) and (ii). Consider the three special vectors,

$$e_1 = (1,0,0), \quad e_2 = (0,1,0), \quad e_3 = (0,0,1).$$

Let us call $f(e_1) = y_1$, and let $y_2 = f(e_2)$, $y_3 = f(e_3)$. And let $y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix}$. If $x = (x_1,x_2,x_3)$, we can write $x = x_1e_1 + x_2e_2 + x_3e_3$. Hence, using properties (i) and (ii), we see that

$$\begin{aligned} f(x) &= f(x_1e_1 + x_2e_2 + x_3e_3) \\ &= f(x_1e_1) + f(x_2e_2) + f(x_3e_3) \\ &= x_1f(e_1) + x_2f(e_2) + x_3f(e_3) \\ &= x_1y_1 + x_2y_2 + x_3y_3 = x \cdot y. \end{aligned}$$

Hence the column vector y represents the linear function f.

Example 1. An office buys three kinds of paper, heavy bond, light bond, and a cheaper quality for intraoffice use. The amounts bought (in reams) are given by the row vector $x = (20,50,70)$. The prices per ream of these types of paper are given (in cents) by the column vector $y = \begin{pmatrix} 160 \\ 140 \\ 120 \end{pmatrix}$. Then $f(x) = x \cdot y = \$186.00$ is the cost of the order. So far y defines a linear function of x. It is customary to give a discount if 100 or more reams are ordered of one item. The new rules for computing the bill define a new function of x, different from f. Let us call the new function by the letter g. Then $g(2x) < 2g(x)$, since the office gets a discount on the light bond and on the cheaper paper. Now we have a function that is not linear. It often happens that a function in science is nearly linear for restricted values of the components, but not even roughly linear outside this range.

Sometimes we assign, not a single quantity to a row vector, but several quantities. Then we say that the vector is *transformed* into another vector. We saw an example (in Section 3) where the row vector giving numbers of houses being built was transformed into a row vector giving amounts of raw materials. We say that the transformation is a *linear transformation* if each component in the resulting vector is a linear function of the given vector.

It follows immediately from the definition of a linear transformation that it can be represented by a set of column vectors, which can

also be written as a matrix. Conversely, every matrix defines a linear transformation of vectors. Finally, it follows from (i) and (ii) that a linear transformation T satisfies the properties $T(kx) = kT(x)$ and $T(x + x') = T(x) + T(x')$, where x and x' are vectors and k is a number.

Example 2. Let us suppose that the population of the United States is divided into five groups according to income. The components of the row vector x are the number of people in each bracket. Say x_1 people have an income of \$100,000 or above, x_2 have incomes between \$20,000 and \$100,000, etc. If we know the average number of cars owned by men in a given income bracket, we can represent these five numbers as a column vector, and we get the number of privately owned cars as a linear function of x. Similarly we could get the number of yachts, privately owned houses, or television sets. Each of these four quantities is a linear function of x (at least approximately) and each is represented by a five-component column vector whose entries are averages. Writing the four vectors together as a rectangular array, we get a 5×4 matrix. This is a linear transformation transforming x into a four-component row vector, whose components are the total number of cars, yachts, houses, and television sets, respectively.

EXERCISES

1. $x = (x_1,x_2,x_3)$. Test each of the following functions of x as to whether it has properties (i) and (ii).

(a) $f(x) = 3x_1 + x_2 - 2x_3$. [*Ans.* Linear.]

(b) $f(x) = x_1x_2x_3$.

(c) $f(x) = \sqrt{(x_1)^2 + (x_2)^2 + (x_3)^2}$. [*Ans.* Not linear.]

(d) $f(x) = x_2$.

2. $x = (x_1,x_2)$. Test each of the following transformations of x into y as to whether it is a linear transformation.

(a) $y_1 = 2x_1 + 3x_2$ and $y_2 = x_1 - x_2$. [*Ans.* Linear.]

(b) $y_1 = x_1 + 2x_2$ and $y_2 = -x_1x_2$. [*Ans.* Not linear.]

(c) $y_1 = x_2$ and $y_2 = -x_1$.

For the linear transformations above, write the matrix representing the transformation.

3. Prove that the function $f(x) = c$, where x is a two-component row vector and c is a constant, is a linear function if and only if $c = 0$.

4. Prove that the function $f(x) = ax_1 + bx_2 + c$, where x is a two-component row rector and a, b, and c are constants, is a linear function if and only if $c = 0$.

5. Prove that the transformation $T(x) = xA + C$, where x is a two-component row vector and A and C are 2×2 matrices, is a linear transformation if and only if $C = 0$.

6. Prove that $f(x) =$ (least component of x) is not a linear function.

7. Let x be a 12-component row vector. Its components are the enrollment figures in twelve mathematics courses. Give an example of

(a) A linear function of x.
[An *Ans.* The total enrollment in all mathematics courses.]
(b) A linear transformation of x.
(c) A nonlinear function of x.

8. Let the components of x be the number of fiction books, the number of nonfiction books, and the number of other publications in a library. For each of the following functions, state whether or not it is a linear function of x.

(a) The total number of publications. [*Ans.* Linear.]
(b) The total number of cards in the catalogue. (Assume that each book has two cards, each other publication has one.)

9. If in (i) and (ii), x is taken as a column vector, then the conditions define a linear function of a column vector. How can we represent such a function? How can we represent a linear transformation of column vectors?

10. Show that the matrix R defined in Section 3 can be thought as a transformation of both row vectors and column vectors.

*10. PERMUTATION MATRICES

In Chapter III we defined a permutation of n objects to be an arrangement of these objects in a definite order. Thus the set $\{a,b,c\}$ has six permutations: *abc*, *acb*, *bac*, *bca*, *cab*, and *cba*. There is a slightly different way of thinking of a permutation. We may think of our set as given originally in a definite order, say *abc*, and then think of a permutation as a *rearrangement* of the set. Thus one permutation changes *abc* into *bac*; i.e., the first element is put into the second spot, the second into the first spot, and the third element is left unchanged. In order to arrive at the same number, $n!$, of permutations as before,

we must consider the "rearrangement" that changes nothing, i.e., the permutation that "changes" *abc* into *abc*. We shall consider our n objects as components of a row vector. A permutation changes the row vector into another having the same components, but possibly in a different order.

A convenient way to describe permutations is by means of certain special matrices. For example, the rearrangement given above can be described by the product

$$(x_1,x_2,x_3)\begin{pmatrix}0 & 1 & 0\\ 1 & 0 & 0\\ 0 & 0 & 1\end{pmatrix} = (x_2,x_1,x_3).$$

In this we do not have to think of the x_i as numbers. They are objects of any sort for which multiplication by 0 and 1 and addition is defined as for numbers. The 3×3 matrix then represents our permutation. It has only 0's and 1's as components, and there is exactly one 1 in each row and in each column.

DEFINITION 1. A *permutation matrix* is a square matrix having exactly one 1 in each row and each column, and having 0's in all other places.

Examples of permutation matrices are shown in Figure 8. Since these matrices are square matrices ($n \times n$), we can speak of the matrix as having *degree n*. Thus Figure 8 shows one matrix of degree 2, two of degree 3, and one of degree 4.

$$A = \begin{pmatrix}0 & 1\\ 1 & 0\end{pmatrix}, \qquad B = \begin{pmatrix}0 & 1 & 0\\ 0 & 0 & 1\\ 1 & 0 & 0\end{pmatrix}, \qquad C = \begin{pmatrix}1 & 0 & 0\\ 0 & 1 & 0\\ 0 & 0 & 1\end{pmatrix},$$

$$D = \begin{pmatrix}0 & 1 & 0 & 0\\ 1 & 0 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}$$

Figure 8

Theorem 1. Every permutation matrix of degree n represents a permutation of n objects, and every such permutation has a unique matrix representation.

Let us consider n objects $x_1, x_2, \ldots, x_n$ which by a permutation are rearranged to give $y_1, y_2, \ldots, y_n$. Here each of the y's is one of the x's, and every x is some y. If it happens that $y_j = x_i$, then the object in the ith position was changed to the jth position. In this case, define $p_{ij} = 1$ and $p_{ik} = 0$ for $k \neq i$. Doing this for every i, we obtain an $n \times n$ permutation matrix P such that

$$(x_1,x_2, \ldots, x_n)P = (y_1,y_2, \ldots, y_n). \tag{1}$$

The fact that no two elements of a single row or a single column of P are 1 (i.e., that P is a permutation matrix) follows from the fact that in a permutation each element appears once and only once in the rearrangement.

On the other hand, if we are given a permutation matrix P, then we can define a permutation by the product (1). The fact that each column of P has exactly one 1 means that each y_j is some x_i. The fact that P has only one 1 in each row means that every x_i appears as only one y_j. Hence the vector $(y_1,y_2, \ldots, y_n)$ does represent a rearrangement of the vector $(x_1,x_2, \ldots, x_n)$, completing the proof of the theorem.

We shall restrict ourselves to the case of $n = 4$ for illustrating the following discussion, but all the results we are about to establish will hold for every n. In Figure 9 we find four examples of permutation matrices of degree 4.

$$I = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}, \qquad J = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \end{pmatrix}$$

$$K = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad L = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Figure 9

We want to study the product of two permutation matrices of degree 4. If $x = (x_1,x_2,x_3,x_4)$, then $xJ = (x_4,x_1,x_3,x_2)$ and $xK = (x_2,x_1,x_4,x_3)$. The former puts the first component into second place, the second component into fourth place, and the fourth component into first place; leaving the third component unchanged. The latter interchanges the first two and the last two. What happens if

we perform the two permutations, one after the other? Let us first consider x_1. In the first transformation it is changed into the second component, while in the second transformation, the second component is changed into the first. Hence x_1 ends up where it started, in first place. The component x_2 is first sent into the number four slot, and then this is changed to number three by the second transformation. Hence x_2 ends up as the third component. Component x_3 is at first not changed, but later changed into component four. Component x_4 is first made into the first component, and in the second transformation it is changed into the second component. Hence, starting with x, after two transformations we end up with (x_1,x_4,x_2,x_3).

Let us now consider the product

$$JK = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix}.$$

The matrix JK is again a permutation matrix, and it is easy to check that it represents precisely the permutation described above.

Theorem 2. The product JK of two permutation matrices of the same degree is again such a permutation matrix. It represents the result of first performing permutation J, then permutation K.

This theorem is very easy to prove in matrix form. We wish to know what $x(JK)$ is. By the associative law (see Section 4) this is the same as $(xJ)K$. But xJ is the result of the J permutation, and $(xJ)K$ is the result of applying the K permutation to xJ. This proves the theorem.

Example. Referring to Figure 9 let us consider the products IJ and JI. We know, of course, that $IJ = JI = J$. Hence Theorem 2 tells us that performing the I permutation followed by the J permutation (or the reverse) will result simply in the J permutation. If we note that the I permutation leaves everything unchanged, this result is obvious.

Let us now consider the product JL, where again J and L are as in Figure 9. The product is equal to I; hence $L = J^{-1}$. By Theorem 2 we know that the permutation J followed by L will result in the permutation I, i.e., in no change at all. Thus we see that $L = J^{-1}$

is a permutation that undoes all changes made by J. We also note a similarity in the structure of J and L; the latter is formed from the former by turning it over its main diagonal (the diagonal slanting from the upper left-hand corner to the lower right-hand corner). In other words L has as its i,jth component what J has as its j,ith component.

DEFINITION 2. The *transpose* A^* of a square matrix A is formed by turning it over its main diagonal; that is, the entries of A^* are given by $a^*_{ij} = a_{ji}$.

Theorem 3. If P is a permutation matrix, then P^* is its inverse; that is, P^* represents the permutation which undoes what the permutation P does.

We must show that P^* undoes what P does; the remainder will follow from the above discussion and Section 6. Let us suppose that $p^*_{ij} = 1$. Then $p_{ji} = 1$; hence the permutation P moves component x_j into position i. But then, because $p^*_{ij} = 1$, the component is moved from position i into position j. Hence x_j ends up in position j, where it started; and this holds for every component. Thus P^* undoes the work of P, which proves the theorem.

DEFINITION 3. A set of objects forms a *group* (with respect to multiplication) if:

(i) The product of two elements of the set is always an element of the set.

(ii) There is in the set an element I, called the identity element, such that for every A in the set, $IA = AI = A$.

(iii) For every A in the set there is an element A^{-1} in the set such that $AA^{-1} = A^{-1}A = I$.

(iv) For every A, B, C in the set, $A(BC) = (AB)C$.

DEFINITION 4. A set of objects form a *commutative group* if in addition to the above four properties they also satisfy:

(v) For every A and B in the set, $AB = BA$.

Theorem 4. The permutation matrices of degree n form a group (with respect to matrix multiplication), but this group is not commutative if $n > 2$.

Proof. Property (i) was shown in Theorem 2. Property (ii) follows from the more general fact that $IM = MI = M$, for every $n \times n$ matrix M. From Theorem 3 we know that A has an inverse, namely $A^{-1} = A^*$. It is easy to show that A^* is again a permutation matrix (see Exercise 1). Hence (iii) follows. And (iv) again follows from the more general theorem that all matrices obey this associative law. (See Section 6.) On the other hand it is easy to show examples, for any $n > 2$, where $AB \neq BA$. (See Exercises 2-3.) This completes the proof.

The group formed by the $n \times n$ permutation matrices is known as the *permutation group of degree n.* Since permutations are used in the study of symmetry, this group is also called the *symmetric group of degree n.*

EXERCISES

1. Prove that the transpose of a permutation matrix is a permutation matrix; i.e., that if A satisfies Definition 1, then so does A^*.

2. Write all permutation matrices of degree 1. Write all permutation matrices of degree 2. Show that these two groups are commutative.

3. For $n > 2$, we can form the matrix A which only interchanges x_1 and x_2, and the matrix B which only interchanges x_1 and x_3. What permutations are performed by AB and by BA? Are these two the same? Use this fact to show that the permutation group of order $n > 2$ is not commutative.

4. Write down the permutation matrices which change (x_1,x_2,x_3,x_4) into:

(a) (x_2,x_3,x_4,x_1).
(b) (x_1,x_3,x_2,x_4).
(c) (x_2,x_3,x_1,x_4).
(d) (x_1,x_2,x_3,x_4).

[*Ans.* (a) $\begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$.]

5. For the following pairs of matrices, find the permutations they represent. In each case show that AB represents the permutation A followed by the permutation B, and that BA represents the permutation B followed by the permutation A.

(a) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $\qquad B = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$,

(b) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$,

(c) $A = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$, $B = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix}$.

[*Ans.* (a) xA is (x_3,x_1,x_2); xB is (x_1,x_3,x_2); xAB is (x_3,x_2,x_1); xBA is (x_2,x_1,x_3).]

6. Prove that the set of *all* 3×3 matrices does *not* form a group (with respect to matrix multiplication).

7. Find the inverses of the six matrices in Exercise 5 by using Theorem 3. Check your answers by multiplying the matrices by their inverses.

[*Ans.* (a) $A^{-1} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}$; $B^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$.]

8. The process of division is usually introduced by saying that b/a is the solution of the equation $ax = b$ (or of $xa = b$).

(a) Prove that in a group the equation $AX = B$ always has a unique solution.

(b) Prove that in a group the equation $XA = B$ always has a unique solution.

(c) Show by means of an example that the two equations need not have the same solution.

9. For the set of numbers {1,2,3,4} we define "multiplication" by means of the following table.

x	1	2	3	4
1	1	2	3	4
2	2	4	1	3
3	3	1	4	2
4	4	3	2	1

(In this table we have neglected all multiples of 5; e.g., $2 \times 4 = 8$, but we neglected the 5 and just kept the remainder 3. Again $3 \times 4 = 12$, but we

ignored the 10, which is a multiple of 5, and kept the remainder 2.) Prove that this set with multiplication so defined forms a commutative group.

10. For the set {1,2,3,4,5,6} write down a multiplication table, ignoring all multiples of 7. (See Exercise 9.) Prove that the result is a commutative group.

11. For the set {1,2,3,4,5} write down a multiplication table, ignoring multiples of 6. (See Exercises 9 and 10.) Prove that the result is *not* a group. Why do 5 and 7 give us groups, but not 6?

12. Write down all permutation matrices of degree 3, and assign letter-names to them. Write a multiplication table for this group. How, from this table alone, can we see that properties (i), (ii), and (iii) hold? How do we see that (v) does not hold?

*11. SUBGROUPS OF PERMUTATION GROUPS

Within a group we sometimes can find smaller groups. Here we shall study some of the subgroups of permutation groups. It will be understood that whenever we speak of a group we have a set with a finite number of elements in mind. In particular this will be assumed for the theorems given below, since some of the theorems are not valid for groups with an infinite number of elements. The concept of a group has important applications for infinite sets, but these do not belong in this book.

DEFINITION 1. If a given set G forms a group, and some subset H of it also forms a group, we call the subset H a *subgroup* of G. If the subset H is a proper subset of G, we speak of a *proper subgroup*.

Theorem 1. If we select any element of a group, the powers of the element form a subgroup which is commutative.

Proof. Select any element A of the given group; we must show that the powers A^n have the properties (i)-(v) given in the last section. The product of two powers is again a power, $A^iA^k = A^{i+k}$; hence (i) holds. Next we observe that the powers cannot be all different, since this would give us infinitely many elements in our group. Hence we must have an equation $A^j = A^k$, with, say, $j > k$. However, this implies that $A^{j-k} = I$. Hence I occurs among the powers of A, say $I = A^m$. Therefore (ii) holds. If $m = 1$ or 2, then A is its own inverse (see Exercise 9). On the other hand, if $m > 2$, then among the powers

we find A^{m-1}, and $AA^{m-1} = A^m = I$, so that A^{m-1} is the inverse of A. This shows that property (iii) holds. The associative law (iv) follows from the fact that all matrices obey this law. Finally, we get commutativity (v) from the fact that $A^iA^k = A^{i+k} = A^{k+i} = A^kA^i$, completing the proof.

DEFINITION 2. A group which consists of the powers of one element A is known as the *cyclic* group *generated* by A.

Thus we know that we can form a cyclic subgroup of a given group by picking any one element A and taking all its powers. The number of elements in this subgroup is called the *order* of A. In the proof above, the order of A is the smallest possible m such that $A^m = I$.

Example 1. The permutation group of degree 4 has $4! = 24$ elements. Let us consider the cyclic subgroup generated by J (see Figure 9). We find that $J^2 = L = J^{-1}$, so that $J^3 = JJ^2 = I$. Thus our cyclic subgroup consists of J, $J^2 = L$, and $J^3 = I$. If we continue to take higher powers, we get $J^4 = J$, $J^5 = L$, $J^6 = I$, etc. The elements are repeated in this fixed cycle. This is the source of the name "cyclic."

Example 2. We can get a larger cyclic subgroup by choosing the matrix M and its powers (see Figure 10); M has order 4; hence $M^{-1} = M^* = M^3$, and $M^4 = I$.

$$M = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}, \qquad M^2 = \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

$$M^3 = \begin{pmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}, \qquad M^4 = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Figure 10

Theorem 2. If in a group we select any subset having property (i), then this subset is a subgroup.

Proof. We must show that the subset also has properties (ii)-(iv). Let A be any element of the subset. By (i), $AA = A^2$ is also in the

subset, and then $AA^2 = A^3$ is in the subset, etc. Hence all powers of A are in the subset. One of these powers is I and one is A^{-1}. Hence we have properties (ii) and (iii). Property (iv) again follows from the fact that all matrices have this property, completing the proof of the theorem.

We now have a practical way of finding subgroups. We select one or more elements of the group, and form all possible products of these, using each one as many times as necessary. If we form all possible products, then the product of any two products will also be on our list, and hence property (i) holds. Then, by Theorem 2, we have a subgroup, which is called the subgroup *generated* by the elements. If we start with a single element, we obtain a cyclic subgroup. Some very interesting subgroups can be generated by two elements.

Example 3. Let us start with J (see Figure 9), and D (see Figure 8), and form the subgroup they generate. First of all we get the powers of J, namely, J and $J^2 = L$ and $J^3 = I$, as was shown in Example 1. Then we have D, and D^2, which is again I. In products formed using both J and D we need consider only J and J^2 and D, since the next higher power is I, and then the powers are repeated. Theoretically we should consider products like $DJDJ^2$ and $JDJDJDJ$, but we can show as follows that such long products give nothing new. First we observe that $DJ = J^2D$, so that in a long product we may always replace DJ by J^2D, and thus put all the J's in front and all the D's at the end. (See Exercise 14.) Therefore the only new products that we need consider are of the form J^aD^b; and since J can occur only to the first or second power and D only to the first power, we arrive at JD and J^2D as the only additional products. Hence our subgroup has six elements: J, J^2, D, I, JD, and J^2D. Since $JD \neq DJ$, the subgroup is *not* commutative.

So far we have found subgroups of 3, 4, and 6 elements. Each of these numbers is a divisor of 24, the total number of elements in the group. It can be shown that the number of elements in a subgroup is always a divisor of the number of elements in the group, but we will not prove that fact here.

Example 4. Let us now form the subgroup generated by D and K. Since $D^2 = I = K^2$, both D and K will occur only to the first

power in a product. Furthermore $DK = KD$; hence the subgroup will have only four elements: I, D, K, DK. This subgroup is commutative. The fact that the subgroup happens to be commutative is a consequence of the following theorem.

Theorem 3. If A and B commute (i.e., $AB = BA$) then any two products formed from A and B also commute. Hence the subgroup generated by A and B is a commutative subgroup.

Proof. Given any product formed from A and B, say $AABBBABAB$, we can make use of the fact that $AB = BA$ to move all the A's up front and all the B's to the end. Hence the product can be written A^iB^j. A second such product can be written A^kB^m. The product of these, $A^iB^jA^kB^m$, can again be rearranged so that all the A's come at the beginning. Hence $(A^iB^j)(A^kB^m) = A^{i+k}B^{j+m} = A^{k+i}B^{m+j} = (A^kB^m)(A^iB^j)$, completing the proof.

We have now found two types of commutative subgroups: (1) cyclic subgroups and (2) subgroups generated by two elements that commute. For the latter it is convenient to have a technique for finding two commuting elements. We will develop one method for finding such pairs.

DEFINITION 3. The *effective set* of a permutation matrix is the set of all those components of the row vector which are changed by the matrix.

For example, D has $\{x_1,x_2\}$ as its effective set, J has $\{x_1,x_2,x_4\}$, K has the set of all four components, and I has the empty set as effective set. K suggests the definition:

DEFINITION 4. A permutation matrix having all the components in its effective set is called a *complete* permutation matrix.

Theorem 4. Two permutation matrices, whose effective sets are disjoint, commute.

Proof. Let A_1 have X_1 as its effective set, and A_2 have X_2, so that $X_1 \cap X_2 = \mathcal{E}$. Then A_1A_2 will make some changes on X_1 and then on X_2. The latter are not affected by the former, since X_1 and X_2 have nothing in common. Thus we get the same result if we perform A_2 followed by A_1.

We now have a simple way of getting a commutative subgroup, other than a cyclic one. Just select any two matrices (other than I) with disjoint effective sets, and form the subgroup that they generate.

EXERCISES

1. Write down the six permutation matrices of degree 3.

2. Form the cyclic subgroup for each of the six matrices in Exercise 1. Are these subgroups all different? What is the order of each matrix?

[*Ans.* Five distinct groups; one of order 1; three of order 2; two of order 3.]

3. Prove that there are no proper subgroups of the permutation group of degree 3, other than those found in Exercise 2.

4. Write the 24 permutation matrices of degree 4.

5. Form the cyclic subgroup for each of the matrices in Exercise 4. How many different ones do you get? What is the order of each matrix?

[*Ans.* 17 distinct groups; one of order 1; nine of order 2; eight of order 3; six of order 4.]

6. Show by an example that the subgroups found in Exercise 5 are not the only proper subgroups of the permutation group of degree 4.

7. Prove the following facts about orders of permutations:

(a) I has order 1.

(b) A permutation which does nothing but interchange one or more pairs of elements has order 2.

(c) Every other permutation has an order greater than 2.

8. Prove that the subgroup generated by A and B is cyclic if and only if one generator is a power of the other.

9. Prove that if a matrix has order 1 or 2, then it is its own inverse.

10. A matrix M is said to be *symmetric* if $m_{ij} = m_{ji}$ for all i and j. Prove that a permutation matrix is symmetric if and only if it has order 1 or 2.

11. Form the subgroup generated by J and K.

[*Ans.* There are 12 elements.]

12. Prove the following facts about effective sets:

(a) I has an effective set of zero elements.

(b) A matrix which simply interchanges two elements has as effective set a set of two elements.

(c) All other matrices have an effective set of at least three elements.

(d) A matrix is complete if and only if the number of elements in its effective set equals its degree.

13. We wish to form a commutative subgroup of the permutation group of degree 4, by means of the method described above. We want to choose two matrices (other than I) with disjoint effective sets, and form the subgroup they generate.

(a) Using the results of Exercise 12, what must the number of elements be in the two effective sets? [*Ans.* 2; 2.]

(b) Choose such a pair of matrices.

(c) Form the subgroup.

14. Prove the following facts about Example 3 above.

(a) $DJ = J^2D$.

(b) From this it follows that $DJ^2 = JD$.

(c) In any product of D's and J's we can put all the J's up front.

15. If A has order m, and m is an even number, then $A^{m/2}$ is its own inverse. Prove this fact. What does this say about an element of order 2?

16. Prove that the cyclic group generated by A^2 is a subgroup of that generated by A. When will this be a proper subgroup?

SUGGESTED READING

Birkhoff, G., and S. MacLane, *A Survey of Modern Algebra*, Macmillan, New York, 1941; rev. ed., 1953, Chaps. VII, VIII.

Weiss, M. J., *Higher Algebra for the Undergraduate*, Wiley, New York, 1949, Chaps. 3, 6, 7.

Johnson, R. E., *First Course in Abstract Algebra*, Prentice-Hall, New York, 1953.

Beaumont, R. A., and R. W. Ball, *Introduction to Modern Algebra and Matrix Theory*, Rinehart, New York, 1954, Chaps. I, II, III, IV.

Chapter VI*

LINEAR PROGRAMMING AND THE THEORY OF GAMES

1. CONVEX SETS

By the *locus* of a linear equation of the form $ax + by = c$ we mean the set of all points whose coordinates (x,y) satisfy the equation. For example, the locus of the equation

(a) $$2x + 3y = 6$$

can be found, by trial and error, to be the straight line plotted in Figure 1. Thus, setting $x = 0$ we get $y = 2$, so that the point (0,2) is on the locus; similarly, $x = 1$ gives $y = \frac{4}{3}$, so that $(1,\frac{4}{3})$ is on the locus; in the same way the point (3,0) is on the locus; etc. We now

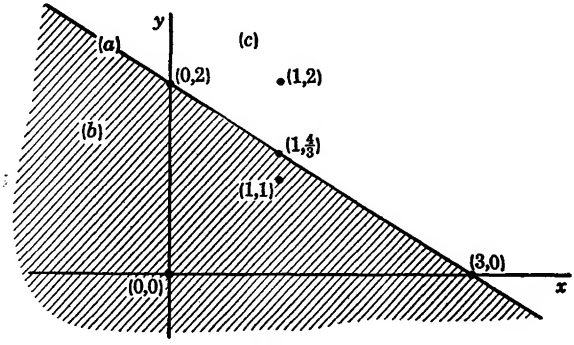


Figure 1

wish to consider the loci of *inequalities* in the variables x and y. Consider

(b) $$2x + 3y < 6$$

as an example. What points (x,y) satisfy this inequality? By trial and error we can find many points on the locus. Thus (1,1) is on it since $2 \cdot 1 + 3 \cdot 1 = 5 < 6$; on the other hand (1,2) is not on the locus since $2 \cdot 1 + 3 \cdot 2 = 8$, which is not less than 6. In between these two points we find $(1,\frac{4}{3})$ which lies on the boundary, i.e., on the locus of (a). We note that by increasing y we went outside the locus; by decreasing y we came into the locus. This holds in general. Given a point on (a), increasing y will give us more than 6, decreasing y gives us less than 6, and hence the latter is on (b). We find that the locus of (b) consists of all points *below* the line (a). This is the shaded area in Figure 1. The area on one side of a straight line is called an *open half plane.*

We can apply exactly the same analysis to

(c) $$2x + 3y > 6$$

to see that its locus is the open half plane above the line (a). This can also be deduced from the fact that the loci (a), (b), and (c) are disjoint and that their union is the entire plane.

If we have an inequality of the form $2x + 3y \leq 6$, its locus will consist of the union of the (a) and (b) loci; hence it consists of an open half plane together with its boundary, which we call a *closed half plane.* The same type of analysis shows that as $ax + by = c$ always has a straight line as its locus, $ax + by < c$ has an open half plane and $ax + by \leq c$ has a closed half plane for a locus.

We will discuss an alternate interpretation of loci, which makes some of the above considerations clearer. An equation or inequality in x and y may be thought of as a statement whose truth or falsity depends on what x and y are, or on what the point (x,y) is. Thus each point of the plane represents one logical possibility, and the entire plane may be thought of as the set of all logical possibilities. Then the so-called locus is simply the truth set of the statement. Since (a), (b), and (c) are a complete set of alternatives (see Chapter I, Section 8) their truth sets are disjoint and have $\mathfrak{U}$ as their union (see Chapter III, Section 1); hence they form a partition of $\mathfrak{U}$. The statement

$2x + 3y \leq 6$ is equivalent to (a) $\vee$ (b), and hence its truth set is the union of the truth sets of (a) and (b).

Suppose now that we consider a *system* of inequalities, that is a set of two or more inequalities, and seek simultaneous solutions to them. For example, consider the system

(d) $x \geq 0$

(e) $y \geq 0$

(f) $2x + 3y \leq 6.$

Here we are asserting three different statements, i.e., we assert their conjunction. Thus the truth set is the intersection of the three individual truth sets. We already know that the locus of (f) is the closed half plane shaded in Figure 1. The locus of (d) is the right-hand closed half plane, while (e) has the upper closed half plane as locus. The intersection of these is the triangle (including the sides) shaded in Figure 2. This area contains all points which satisfy the system of inequalities.

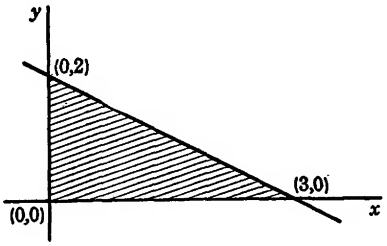


Figure 2

DEFINITION. The intersection of closed half planes is called a *polygonal convex set.*

Theorem. The points which are simultaneous solutions of a system of inequalities of the $\leq$ type form a polygonal convex set.

This theorem follows from the fact that each inequality of the $\leq$ (rather than $<$) type has a closed half plane as its truth set, and that the system has as its truth set the intersection of these half planes.

EXERCISES

1. Draw pictures of the polygonal convex sets which contain the simultaneous solutions to the following systems of inequalities. (Construct the individual closed half planes first, and then take their intersection.)

(a) $x \leq 3.$
$y \leq 2.$
$2x + 3y \geq 0.$

(b) $2x + 3y \geq 6.$
$-x + y \leq 2.$
$x + y \leq 3.$

(c) $2x + 3y \leq 6.$
$-x + y \leq 2.$
$x + y \leq 3.$

(d) $y \geq 0.$
$x \geq 0.$
$x \leq 2.$

(e) $x \leq 2.$
$x \geq -2.$
$y \leq 3.$
$y \geq -3.$

(f) $3x + 2y \leq -6.$
$3x + 2y \leq 6.$

(g) $3x + 2y \geq 6.$
$3x + 2y \leq 6.$

(h) $x - y \geq 0.$
$x + y \leq 0.$

(i) $x \leq 2.$
$x \geq 5.$

(j) $3x + 2y \geq 6.$
$2x + 3y \geq 6.$
$x \geq 0.$
$y \geq 0.$

(k) $2x + y \geq 7.$
$x \leq 0.$
$y \leq 0.$

2. Consider the following sets:
$\mathfrak{U}$ is the whole plane;
A is the half plane which is the locus of $-2x + y < 3.$
B is the half plane which is the locus of $-2x + y > 3.$
C is the half plane which is the locus of $-2x + y \leq 3.$
D is the half plane which is the locus of $-2x + y \geq 3.$
L is the line which is the locus of $-2x + y = 3.$
$\mathcal{E}$ is the empty set.
Show that the following relationships hold among these sets: $\tilde{A} = D$, $\tilde{B} = C$, $\tilde{L} = A \cup B$, $C \cap D = L$, $A \cap B = \mathcal{E}$, $A \cap C = A$, $B \cap D = B$, $A \cup D = \mathfrak{U}$, $B \cup C = \mathfrak{U}$, $A \cup C = C$, $B \cup D = D$, $A \cup L = C$, $B \cup L = D$. Can you find other relationships?

3. Of the polygonal convex sets constructed in Exercise 1, which have a finite area and which have infinite area? What is the boundary of those having finite area?

[*Ans.* (c), (d), (f), (h), and (j) are of infinite area; (g) is a line.]

4. For each of the following half planes give an inequality of which it is the truth set.

(a) The open half plane above the x-axis. [*Ans.* $y > 0.$]

(b) The closed half plane on and above the straight line making angles of 45° with the positive x- and y-axis.

Exercises 5-9 refer to a situation in which a family decides to buy x books and y record albums. The books cost \$4 each, and the albums cost \$3 each.

5. One cannot buy a negative number of books or albums. Write these conditions as inequalities, and draw their truth sets.

6. There are only six books and six albums that they like. Modify the set found in Exercise 5 to take this into account.

7. They are not willing to spend more than \$24 altogether. Modify the set found in Exercise 6.

8. They decide to spend at least twice as much on books as on records. Modify the set of Exercise 7.

9. Finally, they decide that they want to spend \$9 on records. What possibilities are left? [*Ans.* None.]

10. Assume that the following statements are true: Every human being needs at least .02 g of phosphorus per day. Every adult human needs .01 g of calcium, every child (not an infant) needs .03 g of calcium, and every infant needs .04 g of calcium. Plot the amount of phosphorus on the vertical axis and the amount of calcium on the horizontal. Then draw in the convex sets of minimal requirements for adults, infants, and children. State whether or not the following assertions are true.

(a) An adult's needs are fulfilled only if a child's needs are.

(b) If a child's needs are satisfied, then so are an infant's.

(c) A child's needs are satisfied only if an infant's are.

[*Ans.* (a) True. (b) False. (c) False.]

11. Assume that the minimal requirements of human beings are given by the following table:

	Phosphorus	Calcium
Adult	.02	.01
Child	.03	.03
Infant	.01	.02

Plot the amount of phosphorus on the vertical axis and the amount of calcium on the horizontal. Then draw in the convex sets of minimal diet requirements for adults, children (noninfants), and infants. State whether or not the following assertions are true.

(a) If a child's needs are satisfied, so are an adult's.

(b) An infant's needs are satisfied only if a child's needs are.

(c) An adult's needs are satisfied only if an infant's needs are.

(d) Both an adult's and an infant's needs are satisfied only if a child's needs are.
(e) It is possible to satisfy adult needs without satisfying the needs of an infant.

2. MAXIMA AND MINIMA OF LINEAR FUNCTIONS

A polygonal convex set may have either a finite or an infinite area. Both these possibilities are illustrated in Exercise 1 of the previous section. A set of finite area, like that in Figure 2, consists of a polygon together with its interior. It will be convenient to refer to this entire area as a polygon. Hence all polygonal convex sets of finite area are polygons, according to this terminology. Furthermore, each such polygon is a convex set (see Exercise 8 of this section); hence we will refer to them as *convex polygons*.

A polygon with n sides has n corners. For example, the triangle of Figure 2 has the three corners (0,0), (3,0), and (0,2). A corner is formed by the intersection of two lines; hence the corner of one of the polygons is the intersection of the boundaries of two of the half planes.

We can now give an interpretation for the various points of a convex polygon in terms of the system of inequalities. A corner point lies on two boundaries, which means that two of the inequalities are actually equalities. A point on a side, other than a corner point, lies on one boundary and hence one inequality is an equality. An interior point of the polygon must, by a process of elimination, correspond to the case where the inequalities are all strict inequalities, i.e., not only $\leq$ but $<$ holds.

The above description is for the usual case. Should three lines meet in a point, for example, or should there be some redundant inequalities, some obvious modifications would have to be made.

Example 1. Consider the system (d)-(f) of the last section. The corner point (0,0) makes (d) and (e) equalities, i.e., $x = 0$ and $y = 0$. For the corner (0,2) we find that (d) and (f) are equalities, and for (3,0) the pair (e) and (f) turn into equalities. For boundary points other than corners, the side on the y-axis has (d) as an equality, the side on the x-axis has (e) as an equality, and (f) is an equality for the slanting side. For points in the interior of the triangle no equalities hold.

Example 2. Find the polygonal convex set defined by the following system of inequalities:

$$\begin{aligned} 2x + \ y + 9 &\geq 0, \\ -x + 3y + 6 &\geq 0, \\ x + 2y - 3 &\leq 0. \end{aligned}$$

A rough sketch of the three half planes shows that the set is a triangle. Hence we can find the corner points by changing two of the inequalities to equalities and solving simultaneous equations. The first two yield $(-3,-3)$, the first and third $(-7,5)$, while the last two give us $(\frac{21}{5},-\frac{3}{5})$. Hence the polygon is the triangle having these three corners.

Example 3. Let us suppose that in a business problem x and y are quantities we can control, except that there are limitations imposed which can be stated as inequalities. We shall assume that the system of inequalities given in the previous example limits our choice of x and y. Let us assume that a given choice of x and y results in a profit of $x + 2y$ dollars. What is the most and the least profit we can make? We must find the maximum and the minimum value of $x + 2y$ for points (x,y) in the triangle. Let us first try the corners. At $(-3,-3)$ we would have a profit of -9, i.e., a loss of \$9. At $(-7,5)$ we have a profit of \$3, and at $(\frac{21}{5},-\frac{3}{5})$ also a profit of \$3. What can we say about the remainder of the triangle? The last inequality tells us that $x + 2y \leq 3$, hence our profit cannot be more than \$3. If we multiply the first inequality by $\frac{5}{7}$ and the second by $\frac{2}{7}$ and add them, we find that $x + 2y \geq -9$; hence we cannot lose more than \$9. We have thus shown that both the greatest profit and the greatest loss occur at a corner point. We will show that this is true in general.

Given a convex polygon and a function $ax + by$ (thought of as a linear function of points in the plane), we want to show that the maximum and minimum values of the function $ax + by$ always are taken on at a corner point of the polygon. First we will show that the values of $ax + by$ on any line segment lie *between* the values the function has at the two end points (possibly equal to the value at one end point).

We represent the points as row vectors (x,y), and then we see that our function is the linear function represented by the column vector $\begin{pmatrix} a \\ b \end{pmatrix}$. Let the end points of the segment be $p = (x,y)$ and $q = (x',y')$.

We have seen in Chapter V (cf. Figure 4) that the points in between can be represented as $tp + (1 - t)q$, with $0 \leq t \leq 1$. If the values of the function at the points p and q are A and B (assume that $B \leq A$), then at the point in between, the value will be $tA + (1 - t)B$, since the function is linear. This value is $B + (A - B)t$, which is at least B and at most A.

We are now in a position to prove the result illustrated in Example 3.

Theorem. A linear function defined over a convex polygon takes on its maximum (and minimum) value at a corner point of the convex polygon.

The proof of the theorem is illustrated in Figure 3. We shall suppose that at the corner p the function takes on its largest corner value, A, and at the corner q it takes on its smallest corner value, B. Let r be any point of the polygon. Draw a straight line between p

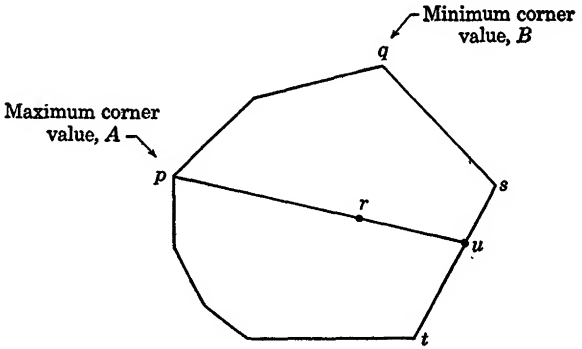


Figure 3

and r and continue it until it cuts the polygon again at a point u lying on an edge of the polygon, say the edge between the corner points s and t. (The line may even cut the edge at one of the points s or t; the analysis remains unchanged.) By hypothesis the value of the function at any corner point must lie between B and A. By the above result the value of the function at u must lie between its values at s and t, and hence must also lie between B and A. Again by the above result the value of the function at r must lie between its values at p

and u, and hence must also lie between B and A. Since r was any point of the polygon our theorem is proved.

Suppose that in place of the linear function $ax + by$ we had considered the function $ax + by + c$. The addition of the constant c merely changes every value of the function, including the maximum and minimum values of the function, by that amount. Hence the analysis of where the maximum and minimum values of the function are taken on is unchanged. Therefore we have the following theorem.

Theorem. The function $ax + by + c$ defined over a convex polygon takes on its maximum (and minimum) value at a corner point of the convex polygon.

The method of finding the maximum or minimum of the function $ax + by + c$ over a convex polygon is then the following: Find the corner points of the set; there will be a finite number of them; substitute the coordinates of each into the function; the largest of the values so obtained will be the maximum of the function and the smallest value will be the minimum of the function. The method is illustrated in Example 3 above.

EXERCISES

1. (a) Draw a picture of the convex polygon obtained in Example 2 above.

(b) Draw a picture of the convex set defined by the inequalities

$$\begin{aligned} 2x + \ y + 9 &\leq 0, \\ -x + 3y + 6 &\leq 0, \\ x + 2y - 3 &\leq 0. \end{aligned}$$

(c) What is the relationship between the two figures?

2. Find the corner points of the convex polygons given in parts (a), (b), and (e) of Exercise 1 following Section 1.

[*Ans.* (a) (3,−2), (3,2), (−3,2); (e) (2,3), (−2,3), (2,−3), (−2,−3).]

3. (a) Show that the three lines whose equations are

$$\begin{aligned} 2x + \ y + 9 &= 0 \\ -x + 3y + 6 &= 0 \\ x + 2y - 3 &= 0 \end{aligned}$$

divide the plane into seven convex regions. Mark these regions with roman numerals I-VII.

(b) For each of the seven regions found in part (a), write a set of three inequalities, having the region as its locus. (*Hint:* Two of these sets of inequalities are considered in Exercise 1.)

(c) There is one more way of putting inequality signs into the three equations given in (a). What is the locus of this last set of inequalities? [*Ans.* The empty set $\mathcal{E}$.]

4. (a) Find the maximum and minimum of the function

$$f(x,y) = -2x + 5y + 17$$

over each of the convex polygons given in parts (a), (b), and (e) of Exercise 1 following Section 1. [*Ans.* (a) 33, 1; (e) 36, −2.]

(b) Find the minimum, when it exists, of the function

$$f(x,y) = 5x + 3y - 6$$

over each of the polygonal convex sets given in parts (h), (i), and (j) of Exercise 1 following Section 1.

[*Ans.* (h) Neither maximum nor minimum; (j) Minimum is 3.]

5. (a) Find the corner points of the convex polygon given by the equations

$$\begin{aligned} 2x + y + 9 &\geq 0, \\ -x + 3y + 6 &\geq 0, \\ x + 2y - 3 &\leq 0, \\ x + y &\leq 0. \end{aligned}$$

(*Hint:* use some of the results of Example 2 in the text above.)

(b) Find the maximum and minimum of the function

$$f(x,y) = 7x + 5y - 3$$

over the convex polygon given in part (a).

[*Ans.* Maximum, 0; minimum, −39.]

6. A convex polygon has the points (−1,0), (3,4), (0,−3), and (1,6) as corner points. Find a set of inequalities which defines the convex polygon having these corner points.

7. Consider the polygonal convex set P defined by the inequalities

$$\begin{aligned} -1 &\leq x \leq 4 \\ 0 &\leq y \leq 6. \end{aligned}$$

Find four different sets of conditions on the constants a and b that the function $f(x,y) = ax + by$ should have its maximum at one and only one

of the four corner points of P. Find conditions that f should have its minimum at each of these points.

[*Ans.* For example, the maximum is at (4,6) if $a > 0$ and $b > 0$.]

8. A set of points is said to be convex if whenever it contains two points it also contains the line segment connecting them. Show that:

(a) If two points are in the truth set of an inequality, then any point on the connecting segment is also in the truth set.

(b) Every polygonal convex set is a convex set in the above-mentioned sense.

9. Give an example of a quadrilateral that is not a convex set.

10. Prove that for any three vectors, p, q, r, the set of all points $ap + bq + cr$ ($a \geq 0, b \geq 0, c \geq 0, a + b + c = 1$) is a convex set. What geometric figure is this locus? [*Ans.* In general, the locus is a triangle.]

3. LINEAR PROGRAMMING PROBLEMS

An important class of practical problems are those which require the determination of the maximum or the minimum of a function of the form $ax + by + c$ of a point defined over a convex set of points. We illustrate these so-called *linear programming problems* by means of the following series of examples.

Example 1. An advertiser wishes to sponsor a television comedy half hour and must decide on the composition of the show. Let x be the number of minutes of commercial time and let y be the number of minutes the comedian appears. By their definition x and y are nonnegative variables. Assume that the advertiser insists that there be at least three minutes of commercials, while the television network insists that the commercial time be limited to at most fifteen minutes. Now the commercial time plus the comedian time must fill up the half hour, i.e., $x + y = 30$. The latter equation can be written as a pair of inequalities $x + y \geq 30$, and $x + y \leq 30$. The inequalities defining our problem now are

$$\begin{aligned} x &\geq 3, \\ x &\leq 15, \\ y &\geq 0, \\ x + y &\geq 30, \\ x + y &\leq 30. \end{aligned}$$

The "polygon" determined by these inequalities is the line segment shown in Figure 4. The corner points of the polygon are (3,27) and (15,15).

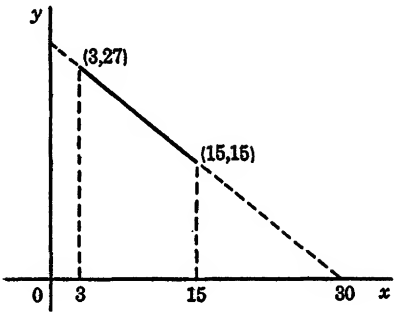


Figure 4

The advertiser has two motives: to minimize the cost of the show, and to maximize the number of people who see it. Suppose the comedian costs \$200 per minute, and the commercials cost \$50 per minute. Then the cost function is

$$C = 50x + 200y.$$

Similarly, suppose that for every minute that the comedian is on the air 70,000 more people will tune in, and for every minute the commercial is on, one more person (e.g., a sponsor) will tune in. Then, if N is the total number of viewers, we have

$$N = x + 70{,}000y.$$

The advertiser now wants to minimize C and maximize N. By substituting in the coordinates of the corner points of the polygon we see that the point (15,15), i.e., 15 minutes of commercials and 15 minutes of comedy, gives the least cost of \$3750 per show. On the other hand, the point (3,27), that is, 3 minutes of commercials and 27 minutes of comedy, gives the maximum number of viewers, namely, 1,890,003. These results are fairly obvious from a common-sense point of view.

Example 2. Suppose now that the comedian (for lack of jokes) refuses to work more than 22 minutes each half hour show. To fill in the remaining time an orchestra is added to the show. Now we have $x + y \leq 30$ as the number of minutes that the commercials and comedian appear, and $30 - x - y$ as the number of minutes that the orchestra plays. Our inequalities now are:

$$\begin{aligned} x &\geq 3, \\ x &\leq 15, \\ y &\geq 0, \\ y &\leq 22, \\ x + y &\leq 30. \end{aligned}$$

The polygon corresponding to these inequalities is shown in Figure 5. Here the set has the five corner points (3,0), (15,0), (15,15), (8,22), and (3,22). Suppose that the band costs \$250 per minute; then the cost function is

$$\begin{aligned} C &= 50x + 200y + 250(30 - x - y) \\ &= 7500 - 200x - 50y. \end{aligned}$$

Here the minimum cost point is (15,15), giving a cost of \$3750 per show. Since $x + y = 30$, we see that the band does not play in this

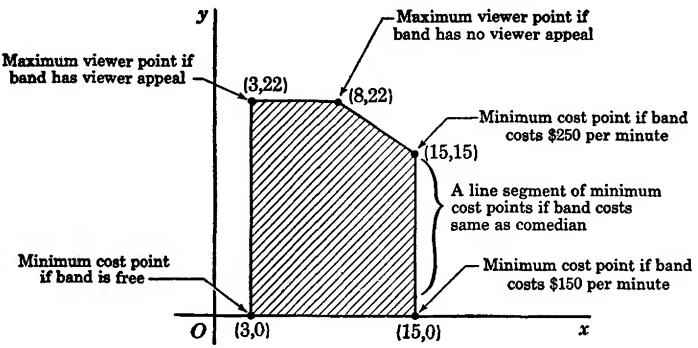


Figure 5

solution. Let us assume first of all that the band has no viewer appeal. Then the N function will be the same as in Example 1. The maximum viewer point will then be (8,22), which gives 1,540,008 viewers. Again $x + y = 30$, so that the band does not play.

Suppose now that the band does have viewer appeal. To be specific, assume that 10,000 more people view the show for each additional minute the band plays. Then our N function becomes

$$\begin{aligned} N &= x + 70{,}000y + 10{,}000(30 - x - y) \\ &= -9{,}999x + 60{,}000y + 300{,}000. \end{aligned}$$

Here the maximum viewer point is (3,22), giving a maximum number of 1,590,003 viewers. Observe that in this solution the band plays five minutes during each show.

Example 3. Our advertiser finds that the financing of the show is becoming difficult and wishes to drop it. To try to induce him to keep

the show the television network reduces the price of the band to \$150 per minute. Then the cost function becomes

$$\begin{aligned} C &= 50x + 200y + 150(30 - x - y) \\ &= -100x + 50y + 4500. \end{aligned}$$

In this case the minimum cost point is (15,0), meaning that the comedian is dropped from the show, which now consists of 15 minutes of commercials and 15 minutes of band music. The minimum cost is now \$3000 per show.

Suppose that this cost is still too high for the advertiser, so that the network offers the band free providing the advertiser still pays for commercials. The cost function is

$$C = 50x + 200y.$$

Here the minimum cost point is (3,0), meaning that the program consists of 27 minutes of band music and 3 minutes of commercials. The minimum cost is \$150 per show.

Example 4. In Example 2 assume that the band and comedian each cost \$200 per minute. Now the cost function is

$$\begin{aligned} C &= 50x + 200y + 200(30 - x - y) \\ &= -150x + 6000. \end{aligned}$$

Here something new happens, since both of the corner points (15,15) and (15,0) yield the minimum cost of \$3750. Observe that the first point has 15 minutes of the comedian and no music, while the second has 15 minutes of music and no comedian. Which of these two solutions should we take? As far as cost goes it doesn't matter, and both are equally good. In fact, if we let the comedian and band divide the 15 minutes arbitrarily, this also is a solution giving minimum cost. Thus all the points $(15,y)$ are possible minimum cost solutions where $0 \leq y \leq 15$.

We have now discovered a general principle. Whenever two corner points each give the same value for our function, the entire connecting line segment also gives this value. This follows from the fact that the value on a line segment is always between the values at the end points (see the last section). Hence, if two corners both give the minimum

(or maximum) value of the function, so does the entire connecting segment.

The above examples show that any one of the corner points or even any point on the whole line segment connecting two of them (and hence any point on the polygon) can be the solution to a linear programming problem depending upon what the facts are and what is desired. But the facts are completely known if we know the values at the corner points.

The exercises below should be worked in the same way. First find the corner points of the convex polygon, then set up the function which is to be maximized or minimized, and then check to see which corner point or points solve the problem.

EXERCISES

1. In Example 2, assume that the comedian and band always cost more than the commercials. Then show that, if the advertiser wishes to minimize cost, the following statements are true.

(a) If the band costs more than the comedian, the band should be dropped.

(b) If the comedian costs more than the band, the comedian should be dropped.

(c) If the comedian and band cost the same, the minimum cost point solution gives 15 minutes of commercials with the remaining 15 minutes being shared in any proportion between the comedian and the band.

2. A well-known nursery rhyme says "Jack Sprat could eat no fat. His wife (call her Jill) could eat no lean. . . ." Suppose Jack wishes to have at least one pound of lean meat per day, while Jill needs at least .4 pound of fat per day. Assume they buy only beef having 10 per cent fat and 90 per cent lean, and pork having 40 per cent fat and 60 per cent lean. Jack and Jill want to fulfill their minimal diet requirements at the lowest possible cost.

(a) Let x be the amount of beef and y the amount of pork which they purchase per day. Construct the convex set of points in the plane representing purchases that fulfill both persons' minimum diet requirements.

(b) Suggest necessary restrictions on the purchases, that will change this set into a convex polygon.

(c) If beef costs \$1 per pound, and pork costs 50 cents per pound, show that the diet of least cost has only pork, and find the minimum cost. [*Ans.* \$.83.]

(d) If beef costs 75 cents and pork costs 50 cents per pound, show that there is a whole line segment of solution points and find the minimum cost. [*Ans.* \$.83.]

(e) If beef and pork each cost \$1 a pound, show that the unique minimal cost diet has both beef and pork. Find the minimum cost. [*Ans.* \$1.40.]

(f) Show that the restriction made in part (b) did not alter the answers given in (c)-(e).

3. In Exercise 2(d) show that for all but one of the minimal cost diets Jill has more than her minimum requirement of fat, while Jack always gets exactly his minimal requirement of lean. Show that all but one of the minimal cost diets contains some beef.

4. In Exercise 2(e) show that Jack and Jill each get exactly their minimal requirements.

5. In Exercise 2, if the price of pork is fixed at \$1 a pound, how low must the price of beef fall before Jack and Jill will eat only beef? [*Ans.* \$.25.]

6. In Exercise 2, suppose that Jack decides to reduce his minimal requirement to 0.6 pound of lean meat per day. How does the convex set change? How do the solutions in 2(c), (d), and (e) change?

7. A poultry farmer raises chickens, ducks, and turkeys and has room for 500 birds on his farm. While he is willing to have a total of 500 birds, he does not want more than 300 ducks on his farm at any one time. Suppose that a chicken costs \$1.50, a duck \$1.00, and a turkey \$4.00 to raise to maturity. Assume that the farmer can sell chickens for \$3.00, ducks for \$2.00, and turkeys for T dollars each. He wants to decide which kind of poultry to raise in order to maximize his profit.

(a) Let x be the number of chickens and y be the number of ducks he will raise. Then $500 - x - y$ is the number of turkeys he raises. What is the convex set of possible values of x and y which satisfy the above restrictions?

(b) Find the expression for the cost of raising x chickens, y ducks, and $(500 - x - y)$ turkeys. Find the expression for the total amount he gets for these birds. Compute the profit which he would make under these circumstances.

(c) If $T = \$6.00$, show that to obtain maximal profit the farmer should raise only turkeys. What is the maximum profit? [*Ans.* \$1000.]

(d) If $T = \$5.00$, show that he should raise only chickens and find his maximum profit. [*Ans.* \$750.]

(e) If $T = \$5.50$, show that he can raise any combination of chickens and turkeys and find his maximum profit. [*Ans.* \$750.]

8. Rework Exercise 7 if the price of chickens drops to \$2.00 and T is (a) \$6.00, (b) \$5.00, (c) \$4.50, and (d) \$4.00.

9. In Exercise 7 show that if the price of turkeys drops below \$5.50, the farmer should raise only chickens. Also show that if the price is above \$5.50, he should raise only turkeys.

10. Let f be a linear function of points (x,y) where (x,y) is a probability vector.

(a) Write the restrictions on x and y.

(b) Find the truth set of this system of conditions.

(c) For what probability vectors could f possibly be a maximum?

[*Ans.* (1,0) or (0,1).]

Exercises 11-20 refer to the following problem. On a chicken farm there are 10 chickens and 32 eggs. In a given time period a chicken can either be used to hatch four eggs or to lay six eggs. The farmer wants to use his available material for two time periods, and then sell all his chickens and eggs. (The product of the first period can be used in the second!) How should he employ his chickens to maximize his income?

11. Let us suppose that x chickens hatch and $10 - x$ lay during the first period. Taking into account the availability of chickens and eggs, what restrictions must be placed on x? (We assume that eggs can be preserved, and hence not all eggs need be used.) [*Ans.* $0 \leq x \leq 8$.]

12. How many chickens will there be at the end of the first period, counting the original chickens plus the newly hatched ones? How many eggs will there be, counting the original ones plus the newly laid ones, less those that were hatched?

13. Let us suppose that in the second period y chickens hatch, and the rest lay. Taking into account the availability of chickens and eggs, what restrictions must be placed on y? How many chickens will lay?

14. Construct the convex polygon determined by the restrictions on x and y. (*Hint:* It has five sides.)

15. How many eggs and how many chickens will there be at the end of the second period? [*Ans.* $152 + 14x - 10y$; $10 + 4x + 4y$.]

16. If eggs sell for 5 cents and chickens for 25 cents, express in terms of x and y the farmer's income. [*Ans.* $1010 + 170x + 50y$.]

17. For what values of x and y is this income largest? How much is the maximum possible income? [*Ans.* $x = 8$, $y = 3$; \$25.20.]

18. For the solution in Exercise 17 trace the number of eggs and chickens at each stage. What characterizes this solution?

[*Ans.* Hatch the eggs as soon as possible.]

19. If eggs sell for 5 cents and chickens for 35 cents, what values of x and y bring in the maximum income? What is it? [*Ans.* 2, 18; \$31.50.]

20. Trace the number of eggs and chickens, and compare with Exercise 18.

4. STRICTLY DETERMINED GAMES

We turn now from linear programming to the theory of games of strategy. Ultimately these two theories can be closely connected, but superficially they are quite different. We can consider a linear programming problem as that of a single person who tries to maximize or minimize a function (of two or more variables) defined over a polygonal convex set of values. In game theory we consider situations in which there are two (or sometimes more) persons, each of whose actions influence, but do not completely determine, the outcome of a single event. The objectives of the players in the game are (usually) different. Game theory provides a solution to such games, based on the principle that each player tries to choose his course of action so that, regardless of what his opponent does, the player can assure himself of a certain amount.

Most recreational games such as tick-tack-toe, checkers, backgammon, chess, poker, bridge, and other card games can be viewed as games of strategy. On the other hand, gambling games such as dice, roulette, etc., are *not* (as usually formulated) games of strategy, since a person playing one of these games is merely "betting against the odds."

The actual games of strategy mentioned above are too complicated, as they stand, to be analyzed completely. We shall instead construct simple examples which, although uninteresting from a player's point of view, do illustrate the theory and which are amenable to computation.

In this section and the next we shall discuss some simple examples of games. The general definition of a matrix game will be given in Section 6.

Example. Consider the following card game: Suppose there are two players, call them R and C (the reason for the use of these letters will be explained later); player R is given a hand consisting of a red 5 and a black 5, while player C is given a black 5 and a red 3. The game that they are to play is the following: At a given signal the players *simultaneously* expose one of their two cards; if the cards

match in color, player R wins the (positive) difference between the numbers on the cards; while if the cards do *not match* in color, player C wins the (positive) difference between the numbers on the cards played. Obviously the strategical decision that each player must make is which of his two cards to play.

A convenient way of representing the game is by means of the matrix shown in Figure 6. (In game theory it is customary to present matrices in this "table" form.) The rows represent the possible choices of player R, and the columns the possible choices of C; hence our use of R and C. The number in position a_{ij} represents the gain

		Player C	
		bk 5	rd 3
Player R	bk 5	0	−2
	rd 5	0	2

Figure 6

of R if R chooses row i and C chooses column j. A positive entry is a payment from C to R, while a negative "gain" for R is a payment from R to C. For example, if R chooses row 1 (plays bk 5) and C chooses column 1 (plays bk 5), then R wins the difference of the two numbers, which is 0. If R chooses row 1 but C chooses column 2 (plays rd 3), then C wins the difference of 2 minus 0, which is indicated by the -2 entry in the matrix. The strategic characteristics of the game are completely described by the matrix.

The game shown in Figure 6 is called a *matrix game.* Any 2×2 matrix can be considered a two-person matrix game by allowing one player to control the rows, the other the columns, and defining the payoffs of the game to be the various matrix entries. In Section 6 we shall see that a matrix of *any* size can in the same way also be considered a matrix game.

How should the players play the matrix game of Figure 6? Player C would like to get the -2 entry in the matrix; however, the only way he could get it would be to play the second column of the matrix, in which case player R would surely choose the second row and C would lose 2 rather than gain 2. On the other hand, if C chooses the first column (i.e., plays bk 5), he assures himself that he will break even regardless of what R does. It is clear that R has nothing to lose

and may possibly gain by choosing the second row, hence he should do so. The knowledge that he will do so reinforces C in his choice of the first column. The optimal procedure for the players is then: R should play rd 5 and C should play bk 5. If they play this way, neither player wins from the other, that is, the game is *fair*.

A command of the form: "Play rd 5," or "Play bk 3," will be called a *strategy*. If player R uses the strategy "Play rd 5" in the game of Figure 6 then, regardless of what C does, R assures himself that he will get *at least* a payoff of zero. Similarly, if C uses the strategy "Play bk 5," then, regardless of what R does, C assures himself of obtaining a payoff of *at most zero*, i.e., a loss of at most zero. Since R cannot, by his own efforts, assure himself of gaining more than zero, and C cannot, by his own efforts, assure himself of losing less than zero, and since these two numbers are the same, we call these *optimal strategies* for the game. Also we call zero the *value* of the game, since it is the outcome of the game if each player uses his optimal strategy.

DEFINITION. We shall say that a 2×2 matrix game is *strictly determined* if the matrix contains an entry, call it v, which is *simultaneously* the *minimum* of the row in which it occurs and the *maximum* of the column in which it occurs. *Optimal strategies* for the players are then the following:

For player R: "Play the row that contains v."
For player C: "Play the column that contains v."

The *value* of the game is v. The game is *fair* if its value is zero.

In Section 6 it will be shown that the strategies here defined are optimal in the sense indicated above, and that v has the property of being the best either player can assure for himself.

The game of Figure 6 is strictly determined, since the 0 entry in the lower left-hand corner of the matrix is the minimum of the second row and the maximum of the first column of that matrix. Observe that the optimal strategies given in the definition above agree with those found above. The value of that game, according to the above definition, is zero; hence it is fair.

The solution of a strictly determined game is particularly easy to find since each player can calculate the other's optimal strategy and hence know what he will do. Not all 2×2 matrix games are so easy to solve, as we shall see in the next section.

In Figure 7 we show three matrix games. The game in Figure 7(a)

0	1
−3	10

(a)

5	2
−7	−4

(b)

0	1
2	0

(c)

Figure 7

is strictly determined and fair, and its optimal strategies are for R to choose the first row and C to choose the first column. The game in Figure 7(b) is strictly determined but not fair, since its value is 2. What are its optimal strategies? Finally, the game in Figure 7(c) is not strictly determined, and the solution of games such as this one will be the subject of the next section.

EXERCISES

1. Determine which of the games given below are strictly determined and which are fair. When the game is strictly determined find optimal strategies for each player.

(a)

0	2
−1	4

(b)

5	0
0	2

(c)

3	1
4	0

(d)

1	−1
−1	1

(e)

3	1
−4	0

(f)

0	4
0	2

(g)

7	0
0	0

(h)

0	0
0	−7

(i)

0	0
0	0

(j)

1	1
1	1

[*Ans.* (a) Strictly determined and fair; R play row 1, C play column 1; (b) nonstrictly determined; (e) strictly determined but not fair; R play row 1, C play column 1; (j) strictly determined but not fair; both players can use any strategy.]

2. In the example suppose that R is given rd 5 and bk 3, and C is given bk 3 and rd 3. Set up the matrix game corresponding to it. Is it strictly determined? Is it fair? Find optimal strategies for each player.

[*Ans.* Yes. Yes. Both play bk 3.]

3. Each of the two players shows one or two fingers (simultaneously) and C pays to R a sum equal to the total number of fingers shown. Write the game matrix. Show that the game is strictly determined, and find the value and optimal strategies.

4. Each of two players shows one or two fingers (simultaneously) and C pays to R an amount equal to the total number of fingers shown, while R pays to C an amount equal to the product of the numbers of fingers shown. Construct the game matrix (the entries will be the net gain of R), and find the value and the optimal strategies.

[*Ans.* $v = 1$, R must show one finger, C may show one or two.]

5. Show that a strictly determined game is fair if and only if there is a zero entry such that both entries in its row are nonnegative and both entries in its column are nonpositive.

6. Consider the game

$$G = \begin{array}{|c|c|} \hline 2 & 5 \\ \hline -1 & a \\ \hline \end{array} .$$

(a) Show that G is strictly determined regardless of the value of a.
(b) Find the value of G. [*Ans.* 2.]
(c) Find optimal strategies for each player.
(d) If $a = 1{,}000{,}000$, obviously R would like to get it as his payoff. Is there any way he can assure himself of obtaining it? What would happen to him if he tried to obtain it?
(e) Show that the value of the game is the most that R can assure for himself.

7. Consider the matrix game

$$G = \begin{array}{|c|c|} \hline a & a \\ \hline c & d \\ \hline \end{array} ;$$

show that G is strictly determined for every set of values for a, c, and d. Show that the same result is true if two entries in a given column are always equal.

8. Find necessary and sufficient conditions that the game

$$G = \begin{array}{|c|c|}\hline a & 0 \\ \hline 0 & b \\ \hline \end{array}$$

should be strictly determined. (*Hint:* These will be expressed in terms of relations among the numbers a and b and the number zero.)

9. Suppose that in the example discussed in the text, player R is given a hand consisting of bk x and rd y, and player C is given bk u and rd v, where x, y, u, and v are real numbers. Suppose that the matrix game which they play is the following:

		Player C: bk u	Player C: rd v
Player R	bk x	$x - u$	$v - x$
	rd y	$u - y$	$y - v$

(a) Show that if $x = u$, $v \geq x$, and $y \geq x$, the game is strictly determined and fair.

(b) Show that if $y = v$, $y \geq x$, and $y \leq u$, the game is strictly determined and fair.

10. Consider a strictly determined 2×2 matrix game G. Suppose u and v are two entries of the matrix such that each is the minimum of the row and the maximum of the column in which it occurs. Show that $u = v$.

5. NONSTRICTLY DETERMINED GAMES

As we saw in the numerical examples of the last section, some matrix games are nonstrictly determined, that is, they have no entry which is simultaneously a row minimum and a column maximum. We can characterize nonstrictly determined 2×2 matrix games as follows:

Theorem. The matrix game

$$G = \begin{array}{|c|c|} \hline a & b \\ \hline c & d \\ \hline \end{array}$$

is nonstrictly determined if and only if one of the following two conditions is satisfied:

(i) $a < b$, $a < c$, $d < b$, and $d < c$.
(ii) $a > b$, $a > c$, $d > b$, and $d > c$.

(These equations mean that the two entries on one diagonal of the matrix must each be greater than each of the two entries on the other diagonal.)

Proof. If either of the conditions (i) or (ii) holds, it is easy to check that no entry of the matrix is simultaneously the minimum of the row and the maximum of the column in which it occurs; hence the game is not strictly determined.

To prove the other half of the theorem, recall that, by Exercise 7 of the last section, if two of the entries in the same row or the same column of G are equal, the game is strictly determined; hence we can assume that no two entries in the same row or the same column are equal.

Suppose now that $a < b$; then $a < c$ or else a is a row minimum and a column maximum; then also $c > d$ or else c is a row minimum and a column maximum; then also $d < b$ or else d is a row minimum and a column maximum. Hence the assumption $a < b$ leads to case (i) above.

In a similar manner the assumption $a > b$ leads to case (ii). This completes the proof of the theorem.

Example 1. Consider the card game of the example in the last section and assume that player R has bk 5 and rd 3 while player C has bk 3 and rd 5. The rules of play are as before. The corresponding matrix game is

		Player C: bk 3	Player C: rd 5
Player R	bk 5	2	0
	rd 3	0	2

which clearly is nonstrictly determined.

Example 2. Consider again Chapter III, Section 2, Example 1. Recall that Jones conceals either a \$1 or a \$2 bill in his hand; Smith guesses 1 or 2, and wins the bill if he guesses its number. The matrix of this game is

		Smith guesses: 1	Smith guesses: 2
Jones chooses	\$1 bill	−1	0
	\$2 bill	0	−2

Again the game is nonstrictly determined.

How should one play a nonstrictly determined game? We must first convince ourselves that no one choice is clearly optimal for either player. In Example 1, R would like to win 2. But if he definitely chooses bk 5, and C finds this out, C can bring about a zero by playing rd 5. If R chooses rd 3, C can bring about a zero by playing bk 3. Similarly, if C's choice is found out by R, then R can win 2. So our first result is that each player must, in some way, prevent the other player from finding out which card he is going to play.

We also note that for a single play of the game there is no difference between the two strategies, as long as one's strategy is not guessed by the opponent. Let us now consider the game being played several times. What should R do? Clearly, he should not play the same card all the time, or C will be able to notice what R is doing, and profit by it. Rather, R should sometimes play one card, and sometimes the other! Our key question then is, "How often should R play each of

his cards?" From the symmetry of the problem we can guess that he should play each card as often as the other, hence each one-half the time. (We will see later that this is, indeed, optimal.) In what order should he do this? For example, should he alternate bk 5 and rd 3? That is dangerous, because if C notices the pattern, he will gain by knowing just what R will do next. Thus we see that R should play bk 5 half the time, but according to some unguessable pattern. The only safe way of doing this is to play it half the time at random. He could, for example, toss a coin (without letting C see it) and play bk 5 if it comes up heads, rd 3 if it comes up tails. Then his opponent cannot guess his decision, since he himself won't know what the decision is. Thus we conclude that a rational way of playing is for each player to *mix* his strategies, selecting sometimes one, sometimes the other; and these strategies should be selected at random, according to certain fixed ratios (probabilities) of selecting each.

By a *mixed strategy* for player R we shall mean a command of the form, "Play row 1 with probability p_1 and play row 2 with probability p_2," where we assume that $p_1 \geq 0$ and $p_2 \geq 0$ and $p_1 + p_2 = 1$. Similarly, a mixed strategy for player C is a command of the form, "Play column 1 with probability q_1 and play column 2 with probability q_2," where $q_1 \geq 0$, $q_2 \geq 0$, and $q_1 + q_2 = 1$. A mixed strategy vector for player R is the probability row vector (p_1,p_2), and a mixed strategy vector for player C is the probability column vector $\begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$.

Examples of mixed strategies are $(\frac{1}{2},\frac{1}{2})$ and $\begin{pmatrix} \frac{1}{5} \\ \frac{4}{5} \end{pmatrix}$. The reader may wonder how a player could actually play one of these strategies. The mixed strategy $(\frac{1}{2},\frac{1}{2})$ is easy to realize since it is simply the coin-flipping strategy described above. The mixed strategy $\begin{pmatrix} \frac{1}{5} \\ \frac{4}{5} \end{pmatrix}$ is more difficult to realize since there is no chance device in common use that gives these probabilities. However, suppose that a pointer is constructed with a card that is $\frac{4}{5}$ shaded and $\frac{1}{5}$ unshaded, as in Figure 8, and C simply spins the pointer (without letting R see it, of course!).

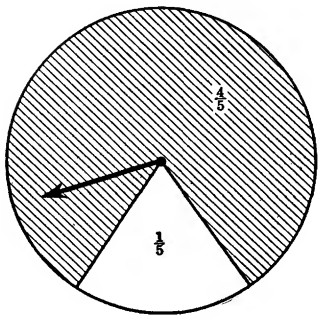


Figure 8

Then, if the pointer stops on the unshaded part he plays the first column, and if it stops on the shaded part, he plays the second column, and thus realizes the desired strategy. By varying the proportion of shaded area on the card other mixed strategies can conveniently be realized.

Consider the nonstrictly determined game

$$G = \begin{array}{|c|c|}\hline a & b \\ \hline c & d \\ \hline \end{array}\,.$$

Having argued, as above, that the players should use mixed strategies in playing a nonstrictly determined game, it is still necessary to decide how to choose an optimal mixed strategy.

DEFINITION. For the nonstrictly determined game G the number v is its value and $p^0 = (p_1^0, p_2^0)$ and $q^0 = (q_1^0, q_2^0)$ are optimal strategies for R and C, respectively, if the following inequalities are satisfied:

$$(1) \qquad p^0 G = (p_1^0, p_2^0)\begin{pmatrix} a & b \\ c & d \end{pmatrix} \geq (v, v);$$

$$(2) \qquad Gq^0 = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} q_1^0 \\ q_2^0 \end{pmatrix} \leq \begin{pmatrix} v \\ v \end{pmatrix}.$$

(If z and w are vectors, the inequality $z \geq w$ means that each component of z is greater than or equal to the corresponding component of w.) The game is *fair* if $v = 0$.

If R chooses a mixed strategy $p = (p_1, p_2)$ and (independently) C chooses a mixed strategy $q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$, then player R obtains the payoff a with probability p_1q_1; he obtains the payoff b with probability p_1q_2; he obtains c with probability p_2q_1; and he obtains d with probability p_2q_2; hence his mathematical expectation (see Chapter IV, Section 12) is then given by the expression

$$ap_1q_1 + bp_1q_2 + cp_2q_1 + dp_2q_2 = pGq.$$

By a similar computation, one can show that player C's expectation is the negative of this expression.

To justify this definition we must show that if v, p^0, q^0 exist for G, each player can guarantee himself an expectation of v. Let q be any strategy for C. Multiplying (1) on the right by q, we get $p^0Gq \geq (v,v)q = v$, which shows that, regardless of how C plays, R can assure himself of an expectation of at least v. Similarly, let p be any strategy vector for R. Multiplying (2) on the left by p, we obtain $pGq^0 \leq p\begin{pmatrix} v \\ v \end{pmatrix} = v$, which shows that, regardless of how R plays, C can assure himself of an expectation of at most v. It is in this sense that p^0 and q^0 are optimal. It follows further that, if both players play optimally, then R's expectation is exactly v and C's expectation is exactly v. (Compare Exercise 11.) Hence we call v the (expected) value of the game.

We must now see whether there are strategies p^0 and q^0 for the game G. While in more complicated games the finding of optimal strategies is a difficult task, for a 2×2 nonstrictly determined game the following formulas provide the solution.

$$p_1^0 = \frac{d - c}{a + d - b - c} \tag{3}$$

$$p_2^0 = \frac{a - b}{a + d - b - c} \tag{4}$$

$$q_1^0 = \frac{d - b}{a + d - b - c} \tag{5}$$

$$q_2^0 = \frac{a - c}{a + d - b - c} \tag{6}$$

$$v = \frac{ad - bc}{a + d - b - c} \tag{7}$$

It is an easy matter to verify (see Exercise 12) that formulas (3)–(7) satisfy conditions (1)–(2). Actually, the inequalities in (1) and (2) become equalities in this simple case, a fact that is not true in general for nonstrictly determined games of larger size.

The denominator in each formula is the difference between the sums of the entries on the two diagonals. Since, for a nonstrictly determined game, the entries on one diagonal must be larger than those on the other, the denominator cannot be zero. The reader will recognize the numerator of v as the determinant.

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Let us use these formulas to solve the examples mentioned earlier.

Example 1 (continued). The solution is easily found by substituting into the above formulas. We obtain $(\frac{1}{2},\frac{1}{2})$ as the optimal strategy for R and $\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\end{pmatrix}$ as the optimal strategy for C. Hence each player should use the coin-flipping strategy for optimal results. The value of the game is plus 1, which means that it is biased in R's favor, and R has an expected gain of 1 per game.

Example 2 (continued). Substitution into the formulas gives $(\frac{2}{3},\frac{1}{3})$ as the optimal strategy for R and $\begin{pmatrix}\frac{2}{3}\\ \frac{1}{3}\end{pmatrix}$ as the optimal strategy for C. The value of the game is $-\frac{2}{3}$, which means that the game is biased in Smith's favor. Smith should then pay $66\frac{2}{3}$ cents to play the game, which answers the question raised in Chapter III.

EXERCISES

1. Find the optimal strategies for each player and the values of the following games:

(a)

1	2
3	4

(b)

1	0
−1	2

(c)

2	3
1	4

(d)

15	3
−1	2

(e)

7	−6
5	8

(f)

3	15
−1	10

[*Ans.* (a) $v = 3$; $(0,1)$; $\begin{pmatrix}1\\ 0\end{pmatrix}$.

(b) $v = \frac{1}{2}$; $(\frac{3}{4}, \frac{1}{4})$; $\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\end{pmatrix}$.

(d) $v = 3$; $(1,0)$; $\begin{pmatrix}0\\ 1\end{pmatrix}$.

(e) $v = \frac{43}{8}$; $(\frac{3}{16}, \frac{13}{16})$; $\begin{pmatrix}\frac{7}{8}\\ \frac{1}{8}\end{pmatrix}$.]

2. Set up the ordinary game of matching pennies as a matrix game. Find its value and optimal strategies. How are the optimal strategies realized in practice by players of this game?

3. A version of two-finger Morra is played as follows: Each player holds up either one or two fingers; if the sum of the number of fingers shown is even, player R gets the sum, and if the sum is odd, player C gets it.

(a) Show that the game matrix is

		Player C	
		1	2
Player R	1	2	−3
	2	−3	4

.

(b) Find optimal strategies for each player and the value of the game. [*Ans.* $(\frac{7}{12}, \frac{5}{12})$; $v = -\frac{1}{12}$.]

4. Rework Exercise 3 if player C gets the even sum and player R gets the odd sum.

5. Consider the following "war" problem: Some attacking bombers are attempting to bomb a city that is protected by fighters. The bombers can each day attack either "high" or "low," the low attack making the bombing more accurate. Similarly, the fighters can each day look for the bombers either "high" or "low." Credit the bombers with six points if they avoid the fighters, and zero if the fighters find them. Also credit the bombers with three extra points for accurate bombing if they fly low.

(a) Set up the game matrix.

(b) Find optimal strategies for each player.

(c) Give instructions to the bomber and fighter commanders so that by flipping coins they can decide what to do.

[*Ans.* (c) The bomber commander should flip one coin to decide whether to go high or low. The fighter commander should flip two coins and then go high if both turn up heads.]

6. Generalize the problem in Exercise 5 by crediting the bombers with x points for avoiding the fighters and y points for flying low. (Assume that x and y are positive.)

(a) Set up the matrix.

(b) If $y \geq x$ show that the game is strictly determined, and find optimal strategies.

(c) If $y < x$ show that the game is nonstrictly determined and find optimal strategies.

(d) Comment on these results, with special attention to the bombers' strategies.

7. If $G = \begin{array}{|c|c|}\hline a & b \\ \hline c & d \\ \hline \end{array}$ is nonstrictly determined, prove that it is fair if and only if

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc = 0.$$

8. In formulas (3)–(7) prove that $p_1 > 0$, $p_2 > 0$, $q_1 > 0$, and $q_2 > 0$. Must v be greater than zero?

9. Utilizing the results of Exercise 7 of the last section, find necessary and sufficient conditions that the game

$$G = \begin{array}{|c|c|}\hline a & 0 \\ \hline 0 & b \\ \hline \end{array}$$

be nonstrictly determined. Find optimal strategies for each player and the value of G, if it is nonstrictly determined.

[*Ans.* a and b must be both positive or both negative. $p_1 = b/(a + b)$; $p_2 = a/(a + b)$; $q_1 = b/(a + b)$; $q_2 = a/(a + b)$; $v = \dfrac{ab}{a + b}$.]

10. Suppose that R is given bk x and rd y while C is given bk u and rd v (where x, y, u, and v stand for positive integers). Let them play the matrix game

	bk u	rd v
bk x	xu	$-xv$
rd y	$-yu$	yv

Show that the game is always nonstrictly determined, and always fair.

11. If G, p^0, q^0, and v are as in the definition, show that $v = p^0 G q^0$.

12. Verify that (3)–(7) satisfy the conditions (1) and (2).

6. MATRIX GAMES

We shall consider a large class of games in this section, and discuss them in considerable generality. Our games are played between two

players, according to strictly specified rules. Each player performs certain actions, as specified by the rules of the game, and then, at the end of the play of the game, one of the players may have to pay a sum of money to the other player. The game may be repeated many times.

During such a game a player may have to make many strategic decisions. By a (pure) *strategy* for one of the players we mean a complete set of rules as to how he should make his decisions. We shall illustrate this in terms of the game of tick-tack-toe (and nearly the same remarks would apply to any game in which the players take turns moving). Let us construct a strategy for the player who moves first. His first decision concerns the opening move. He may choose any one of nine squares, and the strategy must tell him which choice to make. Let us say we tell him to move into the upper left-hand corner. His opponent may answer this in one of eight ways, and the strategy must be prepared for each alternative. It must have eight rules, such as "If he moves into the middle, move into the lower right-hand corner!" For every such move the opponent may respond with one of several alternatives, and the strategy must again have an answering move ready for each of them, etc. Hence the strategy takes into account every conceivable position of the first player, and instructs what move to make in each one.

A strategy may be thought of as a set of instructions to be given to a machine, so that the machine will play the game exactly the way we would have.

We number the strategies of the first player $1, 2, \ldots, m$, and those of the second player $1, 2, \ldots, n$. Since each of the players must play according to one of his strategies, the game may proceed in any one of mn ways, and if each player chooses a definite strategy, the outcome is determined. We may think of giving the two strategies to two machines, and let them work out what happens. Let us suppose that, when the first player chooses strategy i and the second strategy j, the former wins an amount a_{ij}. We arrange these numbers a_{ij} into an $m \times n$ matrix, the *game matrix*. We may then think of the game as consisting of a choice of a row by the first player, and a column by the second player. Hence we see that any game specified by rules may be thought of as a matrix game.

Conversely, every matrix can be considered as a game. An $m \times n$ matrix may be thought of as a game between two players, in which

player R chooses one of the m rows and player C simultaneously chooses one of the n columns. The outcome of the game is that C pays to R an amount equal to the entry of the matrix in the chosen row and column. (A negative entry represents a payment from R to C, as usual.)

In an $m \times n$ matrix game, the player R has m pure strategies, and the player C has n. We have seen in the last section that, in addition, we must consider the mixed strategies of the two players. We extend this concept to $m \times n$ games.

DEFINITION. An m-component row vector p is a *mixed-strategy vector* for R if it is a probability vector; similarly, an n-component column vector q is a mixed-strategy vector for C if it is a probability vector. (Recall from Chapter V that a probability vector is one with nonnegative entries whose sum is 1.) Let V and V' be the vectors

$$V = \underbrace{(v,v,\ldots,v)}_{m \text{ components}} \quad \text{and} \quad V' = \begin{pmatrix} v \\ v \\ \cdot \\ \cdot \\ \cdot \\ v \end{pmatrix} \Big\} n \text{ components}$$

where v is a number. Then v is the *value of the game* and p^0 and q^0 are *optimal strategies* for the players if and only if the following inequalities hold:

$$p^0G \geq V,$$
$$Gq^0 \leq V'.$$

In Sections 4 and 5 we have given several examples of such matrix games together with their solutions. Notice that we have *not* proved that an arbitrary matrix game has a value and optimal strategies for each player; that question will be discussed in the next section.

Theorem. If G is a matrix game which has a value and optimal strategies, then the value of the game is unique.

Proof. Suppose that v and w are two different values for the game G. Let $V = (v,v,\ldots,v)$ and $W = (w,w,\ldots,w)$ be m-component row vectors, and let

$$V' = \begin{pmatrix} v \\ v \\ \cdot \\ \cdot \\ \cdot \\ v \end{pmatrix} \quad \text{and} \quad W' = \begin{pmatrix} w \\ w \\ \cdot \\ \cdot \\ \cdot \\ w \end{pmatrix}$$

be n-component column vectors. Then let p^0 and q^0 be optimal mixed strategy vectors associated with the value v such that

(a) $$p^0G \geq V,$$

(b) $$Gq^0 \leq V'.$$

Similarly, let p^1 and q^1 be optimal mixed strategy vectors associated with the value w such that

(c) $$p^1G \geq W,$$

(d) $$Gq^1 \leq W'.$$

If we now multiply (a) on the right by q^1, we get $p^0Gq^1 \geq Vq^1 = v$. In the same way, multiplying (d) on the left by p^0 gives $p^0Gq^1 \leq w$. The two inequalities just obtained show that $w \geq v$.

Next we multiply (b) on the left by p^1 and (c) on the right by q^0, obtaining $v \geq p^1Gq^0$ and $p^1Gq^0 \geq w$, which together imply that $v \geq w$.

Finally we see that $v \leq w$ and $v \geq w$ imply together that $v = w$, that is, the value of the game is unique.

Theorem. If G is a matrix game with value v and optimal strategies p^0 and q^0, then $v = p^0Gq^0$.

Proof. By definition v, p^0, and q^0 satisfy

$$p^0G \geq V \quad \text{and} \quad Gq^0 \leq V'.$$

Multiplying the first of these inequalities on the right by q^0, we get $p^0Gq^0 \geq v$. Similarly, multiplying the second inequality on the left by p^0, we obtain $p^0Gq^0 \leq v$. These two inequalities together imply that $v = p^0Gq^0$, concluding the proof.

The theorem just proved is important because it permits us to give an interpretation of the *value* of a game as an *expected value* in the sense of probability (see Chapter IV, Section 12). Briefly the interpretation is the following: If the game G is played repeatedly and if

each time it is played player R uses the mixed strategy p^0 and player C uses the mixed strategy q^0, then the value v of G is the expected value of the game for R. The law of large numbers implies that, if the number of plays of G is sufficiently large, then the average value of R's winnings will (with high probability) be arbitrarily close to the value v of the game G.

As an example, let G be the matrix of the game of matching pennies, i.e.,

$$G = \begin{array}{|c|c|} \hline 1 & -1 \\ \hline -1 & 1 \\ \hline \end{array} .$$

As was found in Exercise 2 of the last section, optimal strategies in this game are for R to choose each row with probability $\frac{1}{2}$ and for C to choose each column with probability $\frac{1}{2}$. The value of G is zero. Notice that the only two payoffs that result from a single play of the game are $+1$ and -1, neither of which is equal to the value (zero) of the game. However, if the game is played repeatedly, the average value of R's payoffs will approach zero, which is the value of the game.

Theorem. If G is a game with value v and optimal strategies p^0 and q^0, then v is the largest expectation that R can assure for himself. Similarly, v is the smallest expectation that C can assure for himself.

Proof. Let p be any mixed strategy vector of R and let q^0 be an optimal strategy for C; then multiply the equation $Gq^0 \leq V'$ on the left by p, obtaining $pGq^0 \leq v$. The latter equation shows that, if C plays optimally, the most that R can assure for himself is v. Now let p^0 be optimal for R; then, for every q, $p^0Gq \geq v$, so that R can actually assure himself of an expectation of v. The proof of the other statement of the theorem is similar.

The above theorem gives an intuitive justification to the definition of value and optimal strategies for a game. Thus the value is the "best" that a player can do and optimal strategies are the means of achieving this "best."

DEFINITION. A matrix game G is *strictly determined* if there is an entry g_{ij} in G that is the minimum entry in the ith row and the maxi-

mum entry in the jth column. (By rearranging and renumbering the rows and columns of a strictly determined matrix game G we can assume that g_{11} is an entry that is the minimum of row 1 and the maximum of column 1.)

Theorem. If G is a strictly determined matrix game, arranged as indicated in the definition, the value of the game is $v = g_{11}$. Moreover, optimal strategies for the players are

$$p^0 = (1,0,0, \ldots, 0) \quad \text{and} \quad q^0 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix}.$$

(These optimal strategies simply say that R should choose the row that contains the entry g_{11} (the first row) and C should choose the column that contains the entry g_{11} (the first column). Compare these optimal strategies with those found in Section 4 for strictly determined 2×2 games.)

Proof. Suppose that G is strictly determined and the rows and columns of G are so arranged and numbered that g_{11} is an entry of G that is the minimum of row 1 and the maximum of column 1. Then we set $v = g_{11}$ and let p^0 and q^0 be the strategies as defined in the statement of the theorem. We have

$$\begin{aligned} p^0G &= (g_{11},g_{12}, \ldots, g_{1n}) \\ &\geq (g_{11},g_{11}, \ldots, g_{11}) = V \end{aligned}$$

where we have used the fact that g_{11} was the minimum of the first row. Similarly, using the fact that g_{11} is the maximum of the first column, we have

$$Gq^0 = \begin{pmatrix} g_{11} \\ g_{21} \\ \cdot \\ \cdot \\ \cdot \\ g_{m1} \end{pmatrix} \leq \begin{pmatrix} g_{11} \\ g_{11} \\ \cdot \\ \cdot \\ \cdot \\ g_{11} \end{pmatrix} = V'.$$

From these two inequalities and the definition of a matrix game given above, we conclude that v is the value of the game and p^0 and q^0 are optimal strategies.

Theorem. If g_{11} and g_{ij} are two entries of G that are the minima of the rows and the maxima of the columns in which they occur, then $v = g_{11} = g_{1j} = g_{i1} = g_{ij}$.

Proof. Using the facts that g_{11} and g_{ij} are the minima of the rows and the maxima of the columns in which they occur we see that

$$g_{ij} \geq g_{1j} \geq g_{11}, \qquad g_{ij} \leq g_{i1} \leq g_{11}.$$

(These inequalities are redundant but still true if either $i = 1$ or $j = 1$.) These two sets of inequalities imply that $g_{ij} = g_{1j} = g_{i1} = g_{11} = v$, completing the proof of the theorem.

Example 1. Although we have proved that the value of a game is unique, it may happen that a game has more than one pair of optimal strategies. For example, let G be the game

$$G = \begin{array}{|c|c|c|c|} \hline 1 & 5 & 1 & 7 \\ \hline -2 & 8 & 0 & -9 \\ \hline 1 & 12 & 1 & 3 \\ \hline \end{array}.$$

Then we see that G is strictly determined with value 1, and the optimal strategies are (1,0,0) and (0,0,1) for player R and

$$\begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \end{pmatrix} \quad \text{and} \quad \begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$$

for player C. In the next theorem we shall see that there are still other optimal strategies for this game.

Theorem. If p^0 and p^1 are two optimal strategies for R in a matrix G then the strategy

$$p = ap^0 + (1 - a)p^1,$$

where a is any number satisfying $0 \leq a \leq 1$, is also an optimal strategy for R.

Similarly, if q^0 and q^1 are optimal strategies for C in G, then the strategy

$$q = aq^0 + (1 - a)q^1,$$

where a is any number satisfying $0 \leq a \leq 1$, is also an optimal strategy for C.

Proof. We shall prove the first statement only and leave the second as an exercise (see Exercise 3). It is easy to show that p is a probability vector. By hypothesis we have $p^0G \geq V$ and $p^1G \geq V$. Hence we see that

$$\begin{aligned} pG &= [ap^0 + (1 - a)p^1]G \\ &= ap^0G + (1 - a)p^1G \\ &\geq aV + (1 - a)V = V \end{aligned}$$

which shows that p is also an optimal strategy, completing the proof of the theorem.

This theorem implies that, in Example 1, the strategies of the form $a(1,0,0) + (1 - a)(0,0,1) = (a,0,1-a)$ are optimal for R. It is easy to check that $(\frac{1}{2},0,\frac{1}{2})$ and $(\frac{1}{4},0,\frac{3}{4})$ are optimal and of this form.

EXERCISES

1. Find the value and all optimal strategies for the following games.

(a)

15	2	−3
6	5	7
−7	4	0

[*Ans.* $v = 5$; $(0,1,0)$; $\begin{pmatrix}0\\1\\0\end{pmatrix}$.]

(b)

5	2	−1	−1
1	1	0	1
3	0	−3	7

(c)

0	5	6	−3
1	−1	2	3
1	2	3	4
−1	0	7	5

2. Find the value of and all optimal strategies for the following games.

(a)
5	10	6	5
5	7	8	5
0	5	6	5

(b)
−2	0	−1
−5	7	8

(c)
0	0	1	0
1	0	0	0
1	0	1	0

(d)
3	2	3
6	2	7
5	1	4

[*Ans.* (a) $v = 5$; $(a,1-a,0)$; $\begin{pmatrix} a \\ 0 \\ 0 \\ 1-a \end{pmatrix}$. (d) $v = 2$; $(a,1-a,0)$; $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.]

3. If q^0 and q^1 are optimal strategies for C in the matrix game G, show that the strategy

$$q = aq^0 + (1 - a)q^1,$$

where a is a constant with $0 \leq a \leq 1$, is also optimal in the game G.

4. Verify that the strategies $p^0 = (\frac{1}{3},\frac{1}{3},\frac{1}{3})$ and

$$q^0 = \begin{pmatrix} \frac{1}{3} \\ \frac{1}{3} \\ \frac{1}{3} \end{pmatrix}$$

are optimal in the game G whose matrix is

$$G = \begin{array}{|c|c|c|} \hline 1 & 0 & 0 \\ \hline 0 & 1 & 0 \\ \hline 0 & 0 & 1 \\ \hline \end{array}.$$

What is the value of the game?

5. Generalize the result of Exercise 4 to the game G whose matrix is the $n \times n$ identity matrix.

6. Suppose that player R tries to find C in one of three towns X,Y, and Z.

The distance between X and Y is five miles, the distance between Y and Z is five miles, and the distance between Z and X is ten miles. Assume that R and C can go to one and only one of the three towns and that if they both go to the same town, R "catches" C and otherwise C "escapes." Credit R with ten points if he catches C, and credit C with a number of points equal to the distance he is away from R if he escapes.

(a) Set up the game matrix.
(b) Show that both players have the same optimal strategy, namely, to go to towns X and Z with equal probabilities and to go to town Y with probability $\frac{1}{4}$.
(c) Find the value of the game.

7. A version of five-finger Morra is played as follows: Each player shows from one to five fingers, and the sum is divided by three. If the sum is exactly divisible by three, there is no exchange of payoffs. If there is a remainder of one, player R wins a sum equal to the total number of fingers, while if the remainder is two, player C wins the sum.

(a) Set up the game matrix. (*Hint:* It is 5×5.)
(b) Verify that an optimal strategy for either player is to show one or five fingers with probability $\frac{1}{6}$, to show two or four fingers with probability $\frac{2}{6}$, and to show three fingers with probability $\frac{1}{3}$.
(c) Is the game fair? [*Ans.* Yes.]

8. Consider the following game:

$$G = \begin{array}{|c|c|c|} \hline a & 0 & 0 \\ \hline 0 & b & 0 \\ \hline 0 & 0 & c \\ \hline \end{array} .$$

(a) If a, b, and c are not all of the same sign, show that the game is strictly determined with value zero.
(b) If a, b, and c are all of the same sign, show that the vector

$$\frac{bc}{ab + bc + ca}, \quad \frac{ca}{ab + bc + ca}, \quad \frac{ab}{ab + bc + ca}$$

is an optimal strategy for player R.
(c) Find player C's optimal strategy for case (b).
(d) Find the value of the game for case (b) and show that it is positive if a, b, and c are all positive, and negative if they are all negative.

9. Two players agree to play the following game. The first player will show 1, 2, or 4 fingers. The second player will show 2, 3, or 5 fingers, simul-

taneously. If the sum of the fingers shown is 3, 5, or 9, the first player receives this sum. Otherwise no payment is made.

(a) Set up the game matrix.

(b) Use the results of Exercise 8 to solve the game.

(c) How much should the first player be willing to pay to play the game? [*Ans.* $\frac{45}{256}$.]

10. Consider the (symmetric) game whose matrix is

$$G = \begin{array}{|c|c|c|} \hline 0 & -a & -b \\ \hline a & 0 & -c \\ \hline b & c & 0 \\ \hline \end{array}$$

(a) If a and b are both positive or both negative, show that G is strictly determined.

(b) If b and c are both positive or both negative, show that G is strictly determined.

(c) If $a > 0$, $b < 0$, and $c > 0$, show that an optimal strategy for player R is given by

$$\frac{c}{a-b+c}, \quad \frac{-b}{a-b+c}, \quad \frac{a}{a-b+c}.$$

(d) In part (c) find an optimal strategy for player C.

(e) If $a < 0$, $b > 0$, and $c < 0$ show that the strategy given in (c) is optimal for R. What is an optimal strategy for player C?

(f) Prove that the value of the game is always zero.

11. In a well-known children's game each player says "stone" or "scissors" or "paper." If one says "stone" and the other "scissors," then the former wins a penny. Similarly, "scissors" beats "paper," and "paper" beats "stone." If the two players name the same item, then the game is a tie.

(a) Set up the game matrix.

(b) Use the results of Exercise 10 to solve the game.

12. In Exercise 11 let us suppose that the payments are different in different cases. Suppose that when "stone breaks scissors," the payment is one cent; when "scissors cut paper," the payment is two cents; and when "paper covers stone," the payment is three cents.

(a) Set up the game matrix.

(b) Use the results of Exercise 10 to solve the game.

[*Ans.* $\frac{1}{3}$ "stone," $\frac{1}{2}$ "scissors," $\frac{1}{6}$ "paper"; $v = 0$.]

7. MORE ON MATRIX GAMES: THE FUNDAMENTAL THEOREM

Here we continue the discussion of the basic properties of matrix games. First we show what happens to the game if each entry in the matrix is multiplied by a nonnegative constant or if the same constant is added to each entry in the matrix. Then we discuss the fundamental existence theorem for matrix games.

Theorem. If k is a nonnegative number, i.e., $k \geq 0$, and G is a matrix game with value v, then the game kG is a matrix game with value kv, and every strategy optimal in G is also optimal in kG. (Recall that the matrix kG is obtained from G by multiplying every entry of G by the number k.)

Proof. Let p^0 be an optimal strategy for R in the game G, that is, $p^0G \geq V$. Then we have

$$p^0(kG) = k(p^0G) \geq kV.$$

Similarly, if q^0 is optimal for C in the game G, then

$$(kG)q^0 = k(Gq^0) \leq kV'.$$

These two inequalities show that kv is the value of kG and also that optimal strategies in G are also optimal in the game kG.

It should be observed that it was essential for the proof of this theorem that k be nonnegative, since multiplying an inequality by a *negative* number has the effect of reversing the direction of the inequality sign. The following example shows that the above theorem is false for negative k's.

Example 1. Let $k = -1$ and let G and $(-1)G$ be the matrices

$$G = \begin{array}{|c|c|}\hline 2 & 3 \\ \hline -1 & 0 \\ \hline \end{array} \quad \text{and} \quad (-1)G = \begin{array}{|c|c|}\hline -2 & -3 \\ \hline 1 & 0 \\ \hline \end{array}.$$

Observe that each of these games is strictly determined but that the value of the first game is 2, while the value of the second is 0 (which

is not equal to $(-1)2 = -2$). Moreover, optimal strategies in G are for R to play the first row with probability 1, and for C to play the first column with probability 1, but neither of these strategies is optimal in the game $(-1)G$.

Theorem. Let G be an $m \times n$ matrix game with value v; let E be the $m \times n$ matrix each of whose entries is 1; and let k be *any* constant. Then the game $G + kE$ has value $v + k$, and every strategy optimal in the game G is also optimal in the game $G + kE$. (The game $G + kE$ is obtained from the game G by adding the number k to each entry in G.)

Proof. Let p^0 and q^0 be optimal strategies in G; then $p^0G \geq V$ and $Gq^0 \leq V'$. We have

$$\begin{aligned} p^0(G + kE) &= p^0G + p^0(kE) \\ &= p^0G + k(p^0E) \\ &\geq (v,v, \ldots, v) + (k,k, \ldots, k) \\ &= (v+k,v+k, \ldots, v+k). \end{aligned}$$

Similarly, we have

$$\begin{aligned} (G + kE)q^0 &= Gq^0 + k(Eq^0) \\ &\leq \begin{pmatrix} v \\ v \\ \cdot \\ \cdot \\ \cdot \\ v \end{pmatrix} + \begin{pmatrix} k \\ k \\ \cdot \\ \cdot \\ \cdot \\ k \end{pmatrix} = \begin{pmatrix} v + k \\ v + k \\ \cdot \\ \cdot \\ \cdot \\ v + k \end{pmatrix}. \end{aligned}$$

These inequalities show that the value of the game $G + kE$ is $v + k$ and also show that each strategy optimal in G is optimal in $G + kE$.

Matrix game theory would not be of very great interest unless we knew under what conditions such a game has a solution. The fundamental theorem of game theory is that every matrix game has a solution. The proof of this theorem is too difficult to be included here, but we do discuss its proof for the 2×2 case.

Fundamental theorem. Let G be any $m \times n$ matrix game; then there exists a value v for G and optimal strategies p^0 for player R and

q^0 for player C. In other words, every matrix game possesses a solution.

Proof for 2 × 2 matrices. If G is strictly determined, the value and optimal strategies were found in Section 4. If G is not strictly determined, formulas (3) through (7) of Section 5 give the optimal strategies and value for G. Since G must be either strictly determined or nonstrictly determined we have covered all cases.

EXERCISES

1. Find the values of the games kG and $G + kE$ for each of the games G whose matrices are given in Exercise 1 of Section 6, if k takes on the values 3, 0, and -2.

2. If G is any matrix game and $k = 0$ find all optimal strategies for each player in the game kG. [*Ans.* Any strategy is optimal.]

3. If G is any matrix game and $k > 0$, show that every strategy optimal in kG is also optimal in G. (*Hint:* Multiply by $1/k$.)

4. If G is any matrix game and k is any constant, show that every strategy optimal in the game $G + kE$ is also optimal in the game G.

5. Suppose that before C and R play a matrix game G, player C gives to player R a payment of k dollars. In this case we shall say C has made a *side payment* of k to R. (If k is negative, then, as usual, this will be a side payment of R to C.)

(a) If C has made a side payment of k to R before playing the game G, show that the game they actually play is $G + kE$.

(b) If v is the value of the game G, find the value of the game $G - vE$.

(c) Using the results of (a) and (b), show that any matrix game G with value v can be made into a fair matrix game by requiring that C make a side payment of $-v$ to R before they play the game G.

6. Show that any matrix game G can be made into a fair matrix game, with each entry in the matrix lying between -1 and 1, by adding the same number to each entry in the matrix and by multiplying each entry by a positive number.

7. Show that the sets of optimal strategies for each player are unchanged by the transformation suggested in Exercise 6. How does the value of the game change?

8. Consider the matrix game:

a	b	b
b	a	b
b	b	a

where $a > b$.

(a) Show that this can be obtained from the identity matrix by multiplying it by a suitable number, and then adding bE.

(b) Use the results of Section 6, Exercise 4, to solve the game.

[*Ans.* $v = a/3 + 2b/3$.]

9. Suppose that the entries of a matrix game are rewritten in new units (e.g., dollars instead of cents). Show that the monetary value of the game has not changed.

10. Consider the game of matching pennies whose matrix is

1	−1
−1	1

.

If the entries of the matrix represent gains or losses of one penny, would you be willing to play the game at least once? If the entries represent gains or losses of one dollar would you be willing to play the game at least once? If they represent gains or losses of one million dollars would you play the game at least once? In each of these cases show that the value is zero and optimal strategies are the same. Discuss the practical application of the theory of games in the light of this example.

8. $2 \times n$ AND $m \times 2$ MATRIX GAMES

After the 2×2 games, the simplest matrix games are the $2 \times n$ and $m \times 2$ games, i.e., where one player has only two strategies. Here we discuss the solution of such games.

Example 1. Suppose that Jones conceals one of the following 4 bills in his hand: a \$1 or a \$2 United States bill or a \$1 or a \$2 Canadian bill. Smith guesses either "United States" or "Canadian" and gets the bill if his guess is correct. The matrix of the game is the following:

			Smith Guesses U.S.	Smith Guesses Can.
Jones Chooses	U.S.	\$1	−1	0
		\$2	−2	0
	Can.	\$1	0	−1
		\$2	0	−2

It is obvious that Jones should always choose the \$1 bill of either country rather than the \$2 bill, since by doing so he may cut his losses and will never increase them. This can be observed in the matrix above, since every entry in the second row is less than or equal to the corresponding entry in the first row, and every entry in the fourth row is less than or equal to the corresponding entry in the third row. In effect we can eliminate the second and fourth rows and reduce the game to the following 2×2 matrix game:

			Smith Guesses U.S.	Smith Guesses Can.
Jones Chooses	U.S.	\$1	−1	0
	Can.	\$1	0	−1

The new matrix game is nonstrictly determined with optimal strategies $(\frac{1}{2},\frac{1}{2})$ for Jones and $\begin{pmatrix}\frac{1}{2}\\ \frac{1}{2}\end{pmatrix}$ for Smith. The value of the game is $-\frac{1}{2}$, which means that Smith should pay 50 cents to play it.

DEFINITION. Let A be an $m \times n$ matrix game. We shall say that row i *majorizes* row h if every entry in row i is as large as or larger than the corresponding entry in row h. Similarly, we shall say that column j *minorizes* column k if every entry in column j is as small as or smaller than the corresponding entry in column h.

Any majorized row or minorized column can be omitted from the matrix game without affecting its solution. In the original matrix of

Example 1 above, we see that row 1 majorizes row 2, and also that row 3 majorizes row 4.

Example 2. Consider again the card game of Section 4, this time giving R a bk 5 and rd 3, while C receives a bk 6 and a bk 5 and a rd 4 and a rd 5. The matrix of the game is

	bk 6	bk 5	rd 4	rd 5
bk 5	1	0	−1	0
rd 3	−3	−2	1	2

.

Observe that column 3 minorizes column 4; that is, C should never play rd 5. Thus our game can be reduced to the following 2×3 game:

	bk 6	bk 5	rd 4
bk 5	1	0	−1
rd 3	−3	−2	1

.

No further rows or columns can be omitted; hence we must introduce a new technique for the solution of this game. It can be shown (though we will not attempt to do so) that, in order to solve a $2 \times n$ or $m \times 2$ game, it is sufficient to look at a number of 2×2 games. These are obtained by striking out columns (or rows) of the original game, till it is reduced to a 2×2 game; these are *derived games* of the original game. It can then be shown that the optimal strategy of each player is optimal in one of the derived games, and that the value of the game is the value of one of the derived games. Hence we need only solve all the 2×2 derived games, and try out each strategy of each player, and each value, to see which are the optimal strategies and the value of the whole game.

In the above 2×3 game we have three derived games:

	bk 6	bk 5
bk 5	1	0
rd 3	−3	−2

	bk 6	rd 4
bk 5	1	−1
rd 3	−3	1

	bk 5	rd 4
bk 5	0	−1
rd 3	−2	1

The first game is strictly determined, the others are not. The optimal strategies of player R in each of these derived games are: (1,0), $(\frac{2}{3},\frac{1}{3})$, and $(\frac{3}{4},\frac{1}{4})$, respectively. The optimal strategies for C are

$$\begin{pmatrix}0\\1\end{pmatrix}, \quad \begin{pmatrix}\frac{1}{3}\\\frac{2}{3}\end{pmatrix}, \quad \begin{pmatrix}\frac{1}{2}\\\frac{1}{2}\end{pmatrix},$$

respectively. The values of the games are 0, $-\frac{1}{3}$, $-\frac{1}{2}$. If we apply the three R-strategies to the original game we note that with the first strategy R may lose as much as 1, with the second he may lose $\frac{2}{3}$, while with the third he cannot lose more than $\frac{1}{2}$. Hence $(\frac{3}{4},\frac{1}{4})$ is optimal for him, and $-\frac{1}{2}$ is the value of the game. In order to apply the C-strategies from a 2×2 derived game to the whole game, we must first extend them to four-component strategies. They then become

$$\begin{pmatrix}0\\1\\0\\0\end{pmatrix}, \quad \begin{pmatrix}\frac{1}{3}\\0\\\frac{2}{3}\\0\end{pmatrix}, \quad \begin{pmatrix}0\\\frac{1}{2}\\\frac{1}{2}\\0\end{pmatrix}.$$

It is easy to verify the last one is optimal for C.

The solution of another $2 \times n$ or $m \times 2$ matrix game should be carried out in a similar manner. First check to see whether or not the game is strictly determined. If it is not strictly determined, eliminate all majorized rows and minorized columns. Then solve all possible 2×2 derived games obtained by striking out one or more rows or columns. The value of the original game will be found as a value of one of these 2×2 games, and the optimal strategies of the original game will be found among the optimal strategies of the derived games (which may have to be extended by the addition of zeros).

Example 3. A numerical example of a 3×2 game is

6	-1
0	2
4	3

.

Here the game is strictly determined, since the entry 3 is the minimum of its row and the maximum of its column. The value of the game is 3, and optimal strategies are $p^0 = (0,0,1)$ and $q^0 = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$.

Example 4. Another numerical example is

1	−1	2	−3
−1	1	0	1

.

Here the fourth column minorizes the second, and the first column minorizes the third. The game is then reduced to

1	−3
−1	1

whose value is $-\frac{1}{3}$, and optimal strategies are $p^0 = (\frac{1}{3},\frac{2}{3})$ and $q^0 = \begin{pmatrix} \frac{2}{3} \\ \frac{1}{3} \end{pmatrix}$; the latter strategy extends to the strategy

$$\begin{pmatrix} \frac{2}{3} \\ 0 \\ 0 \\ \frac{1}{3} \end{pmatrix}$$

which is optimal in the original game.

EXERCISES

1. Solve the following games:

(a)

3	0
−2	3
7	5

[*Ans.* $v = 5$; $(0,0,1)$; $\begin{pmatrix} 0 \\ 1 \end{pmatrix}$.]

(b)

10	5	4	6
18	3	3	4

(c)

1	0	2
0	3	2

[*Ans.* $v = \frac{3}{4}$; $(\frac{3}{4}, \frac{1}{4})$; $\begin{pmatrix} \frac{3}{4} \\ \frac{1}{4} \\ 0 \end{pmatrix}$.]

(d)

0 2
1 3
−1 0
2 0

(e)

1	2	3
4	2	1

[*An ans.* $v = 2$; $(\frac{3}{5}, \frac{2}{5})$; $\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$.]

(f)

1	0	1	1	2
0	−1	−2	−3	−10

2. Solve the following games:

(a)

0	15
8	0
−10	20
10	12

(b)

−1	−2	0	−3	−4
−2	1	0	2	5

(c)

−1	5	−1	−2	8	10
3	−6	0	8	−9	−8

[*An ans.* $v = -\frac{1}{2}$; $(\frac{1}{2},\frac{1}{2})$; $\begin{pmatrix} 0 \\ \frac{1}{12} \\ \frac{11}{12} \\ 0 \\ 0 \\ 0 \end{pmatrix}$.]

3. Solve the game

1	2	3
3	2	1

.

Since there is more than one optimal strategy for C, find a range of optimal strategies for him. (See Section 6, Exercise 3.)

4. In the card game of Example 2 suppose that R has bk 9, bk 5, rd 7 and rd 3, while C has bk 8 and rd 4. Set up and solve the corresponding matrix game.

[*Ans.* $v = 1$; R shows bk 5 and rd 7 each with probability $\frac{1}{2}$; C shows each of his cards with probability $\frac{1}{2}$.]

5. Suppose that Jones conceals in his hand one, two, three, or four silver dollars and Smith guesses "even" or "odd." If Smith's guess is correct, he wins the amount which Jones holds, otherwise he must pay Jones this amount. Set up the corresponding matrix game and find an optimal strategy for each player in which he puts positive weight on all his (pure) strategies. Is the game fair?

6. Consider the following game: Player R announces "one" or "two"; then, independently of each other, both players write down one of these two numbers. If the sum of the three numbers so obtained is odd, C pays R the odd sum in dollars; if the sum of the three numbers is even, R pays C the even sum in dollars.

(a) What are the strategies of R? (*Hint:* He has four strategies.)

(b) What are the strategies of C? (*Hint:* We must consider what C does after "one" is announced after a "two." Hence he has four strategies.)

(c) Write down the matrix for the game.

(d) Restrict player R to announcing "two," and allow for C only those strategies where his number does not depend on the announced number. Solve the resulting 2×2 game.
(e) Extend the above mixed strategies to the original game, and show that they are optimal.
(f) Is the game favorable to R? If so, by how much?

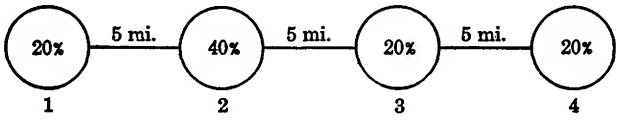
7. Answer the same questions as in Exercise 6 if R gets the even sum and C gets the odd sum (except that, in part (d) restrict R to announce "one"). Which game is more favorable for R? Could you have predicted this without the use of game theory?

8. Rework the five-finger Morra game of Section 6, Exercise 7, with the following payoffs: If the sum of the number of fingers is even, R gets one, while if the sum is odd, C gets one. Suppose that each player shows only one or two fingers. Show that the resulting game is like matching pennies. Show that the optimal strategies for this game, when extended, are optimal in the whole game.

9. A version of three-finger Morra is played as follows: Each player shows from one to three fingers; R always pays C an amount equal to the number of fingers that C shows; if C shows exactly one more or two fewer fingers than R, then C pays R a positive amount x (where x is independent of the number of fingers shown).
(a) Set up the game matrix for arbitrary x's.
(b) If $x = \frac{1}{2}$, show that the game is strictly determined. Find the value. [*Ans.* $v = -\frac{5}{2}$.]
(c) If $x = 2$, show that there is a pair of optimal strategies in which the first player shows one or two fingers and the second player shows two or three fingers. (*Hint:* Solve a 2×2 derived game.) Find the value. [*Ans.* $v = -\frac{3}{2}$.]
(d) If $x = 6$, show that an optimal strategy for R is to use the mixed strategy $(\frac{1}{3},\frac{1}{2},\frac{1}{6})$. Show that the optimal mixed strategy for C is to choose his three strategies each with probability $\frac{1}{3}$. Find the value of the game.

10. Another version of three-finger Morra goes as follows: Each player shows from one to three fingers; if the sum of the number of fingers is even, then R gets an amount equal to the number of fingers that C shows: if the sum is odd, C gets an amount equal to the number of fingers that R shows.
(a) Set up the game matrix.
(b) Reduce the game to a 2×2 matrix game.
(c) Find optimal strategies for each player and show that the game is fair.

11. Two companies, one large and one small, manufacturing the same product, wish to build a new store in one of four towns located on a given highway. If we regard the total population of the four towns as 100 per cent, the distribution of population and distances between towns are as shown:



Assume that if the large company's store is nearer a town it will capture 80 per cent of the business, if both stores are equally distant, then the large company will capture 60 per cent of the business, and if the small store is nearer, then the large company will capture 40 per cent of the business.

(a) Set up the matrix of the game.
(b) Test for majorized rows and minorized columns.
(c) Find optimal strategies and value for the game and interpret your results.

[*Ans.* Both companies should locate in town 2; the large company captures 60 per cent of the business.]

12. Rework Exercise 11 if the per cent of business captured by the large company is 90, 75, and 60, respectively.

13. We have stated without proof that any $2 \times n$ game can be solved by considering only its 2×2 derived games. Verify that this is the case for a game of the form

	C		
R	a	0	1
	0	b	1

.

(a) Show that if $a \leq 1$ or $b \leq 1$, then column 3 is minorized. Hence solve the game.
(b) If $a > 1$ and $b > 1$, solve the three 2×2 derived games. (*Hint:* Two of them are strictly determined.)
(c) If $a > 1$, $b > 1$, but $ab < a + b$, then show that the strategies of the nonstrictly determined derived game are optimal for both players.
(d) If $ab \geq a + b$, then show that R has as optimal strategy the same strategy as in part (c), but C has a pure strategy as optimal strategy.

(e) Using the previous results, show that the value of the game is always the smallest of the values of the three derived games.

9. SIMPLIFIED POKER

In order to illustrate the procedure of translating a game specified by rules into a matrix game, we shall carry it out for a simplification of a well-known game. The example that we are about to discuss is a simplification (by A. W. Tucker) of the poker game discussed on pp. 211–219 in the book *The Theory of Games and Economic Behavior*, by John von Neumann and Oskar Morgenstern.

The deck that is used in simplified poker has only two types of cards, in equal numbers, which we shall call "high" and "low." For example, an ordinary bridge deck could be used with red cards high and black cards low. Each player "antes" an amount a of money and is dealt a single card which is his "hand." By a "deal" we shall mean a pair of cards, the first being given to player R and the second to player C. Thus the deal (H,H) means that each player obtains a high card. There are then four possible deals, namely,

$$(H,H), \quad (H,L), \quad (L,H), \quad (L,L).$$

Ignoring minor errors (see Exercise 1), if the number of cards in the deck is large, each of these deals is "equally likely," that is, the probability of getting a specific one of these deals is $\frac{1}{4}$.

After the deal, player R has the first move and has two alternatives, namely, to "see," or to "raise" by adding an amount b to the pot. If R elects to see, the higher hand wins the pot or equal hands split the pot equally. If R elects to raise, then C has two alternatives, to "fold," or to "call" by adding the amount b to the pot. If C folds, player R wins the pot (without revealing his hand). If C calls, then the higher hand wins the pot or equal hands split the pot equally. These are all the rules.

A pure strategy for a player is a command that tells him exactly what to do in every conceivable situation that can arise in the game. An example of a pure strategy for R is the following: "Raise if you get a high card, and see if you get a low card." We can abbreviate this strategy to simply raise-see. It is easy to see that R has four pure strategies, namely, raise-raise, raise-see, see-raise, and see-see.

In the same manner, C has four pure strategies, fold-fold, fold-call, call-fold, call-call.

Given a choice of a pure strategy for each player, there are exactly four ways the play of the game can proceed, depending on which of the four deals occurs. For example, suppose that R has chosen the see-raise strategy, and C has chosen the fold-fold strategy. If the deal is (H,H), then R sees, and they split the pot, so neither wins; if the deal is (H,L), then R sees and wins the pot, giving him a; if the deal is (L,H), then R raises and C folds, so that R wins a; and if the deal is (L,L), then R raises and C folds, so that R wins a. Since the probabilities of each of these deals is $\frac{1}{4}$, the expected value of R's gain is $3a/4$. Let us compute another expected value, namely, suppose that R uses see-raise and C uses call-fold. Then, if the deal is (H,H), R sees and wins nothing; if the deal is (H,L), then R sees and wins a; if the deal is (L,H), then R raises, C calls, and C wins $a + b$; and if the deal is (L,L), then R raises, C folds, and R wins a. The expected value for R here is $(a - b)/4$.

Continuing in this manner we can compute the expected outcome for each of the 16 possible choice of pairs of strategies. The payoff matrix so obtained is given below.

High		fold	fold	call	call
	Low	fold	call	fold	call
see	see	0	0	0	0
see	raise	$\frac{3a}{4}$	$\frac{2a}{4}$	$\frac{a-b}{4}$	$\frac{-b}{4}$
raise	see	$\frac{a}{4}$	$\frac{a+b}{4}$	0	$\frac{b}{4}$
raise	raise	$\frac{4a}{4}$	$\frac{3a+b}{4}$	$\frac{a-b}{4}$	0

The reader should observe that we have just completed the translation of a game specified by rules into a matrix game.

Since a and b are positive numbers, we see that, in the matrix above,

the fourth row majorizes the second, and the third row majorizes the first. Similarly, the third column minorizes the first and second columns. We can reduce the 4×4 matrix to the following 2×2 matrix:

			Conservative	Bluffing
	High		call	call
		Low	fold	call
Conservative	raise	see	0	$\frac{b}{4}$
Bluffing	raise	raise	$\frac{a-b}{4}$	0

Notice that we have labeled the raise-see strategy as "conservative" for R, since it seems sensible to raise when he has a high card and to see when he has a low one. The strategy raise-raise which says, raise even if you have a low card, we have labeled "bluffing," since it corresponds to the ordinary notion of bluffing. In the same manner we have labeled the call-fold strategy "conservative," and the call-call strategy "bluffing," for player C.

Example 1. Suppose $a = 4$ and $b = 8$. Then the matrix becomes

	Conservative	Bluffing
Conservative	0	2
Bluffing	−1	0

.

Here the game is strictly determined and fair, and optimal strategies are for each player to play conservatively.

Example 2. Suppose $a = 8$ and $b = 4$. Then the matrix becomes

	Conservative	Bluffing
Conservative	0	1
Bluffing	1	0

.

Here the value of the game is $\frac{1}{2}$, meaning that it is biased in favor of R. Optimal strategies are for each player to bluff with probability $\frac{1}{2}$ and to play conservatively with probability $\frac{1}{2}$.

Here we have one of the most interesting results of game theory, since it turns out that, as part of an optimal strategy, one *should* actually bluff part of the time.

EXERCISES

1. Suppose that the simplified poker game is played with an ordinary bridge deck where red is "high" and black is "low." Compute to four decimal places the conditional probability of drawing a red card, given that one red card has already been drawn. From this, discuss the accuracy of the assumption that the four deals are equally likely. How could the accuracy of the assumption be improved?

2. Substitute $a = 4$ and $b = 8$ into the 4×4 matrix above, and reduce it by majorizations and minorizations to a 2×2 matrix game. Is it the one considered in Example 1 above? Do the same for $a = 8$ and $b = 4$ and Example 2.

3. If $a \leq b$, show that the simplified poker game is strictly determined and fair. Show that both players' optimal strategy is to play conservatively.

4. If $a > b$, show that the simplified poker game is biased in favor of R. Show that, to play optimally, each player must bluff with positive probability, and find the optimal strategies.

5. If $a > b$, discuss ways of making the game fair.

6. When $b \geq a$, show that the optimal strategy of player R is not unique. Show that although he has two "optimal" strategies, the raise-see strategy is in a sense better than the other.

7. Show that in the case $a = 8, b = 4$, the strategy of R can be interpreted as follows: "On a high card always raise, on a low card raise with probability $\frac{1}{2}$." Reinterpret C's mixed strategy similarly.

The remaining exercises concern a variant of the simplified poker game. Real poker is characterized by the fact that there are very many poor hands, and very few good ones. We can make the above model of poker more realistic by making the draw of a low card more probable than that of a high card. Let us say that the probability of drawing a high card is only $\frac{1}{5}$. The rules of the game remain as in the text.

8. Calculate the probabilities of (H,H), (H,L), (L,H), and (L,L) deals.

9. The strategies of the two players are as in the text, hence we will get a similar 4×4 game matrix. Calculate the see-raise *vs.* fold-fold entry of the

matrix, just as in the text, but using the results of Exercise 8. Do the same for the see-raise *vs.* call-fold entry. [*Ans.* $24a/25$; $(16a - 4b)/25$.]

10. Fill in the remaining matrix entries.

11. Show that two rows are majorized, and that two columns are minorized.

12. Show that the resulting 2×2 game is strictly determined if and only if $b \geq 4a$. What is the value of the game in these cases?

13. Let $a = 4$, $b = 8$, as in the text, and solve the game. Compare your solution with that in the text.

[*Ans.* Each player should bluff half the time; $v = \frac{16}{25}$; in the previous version there was no bluffing in this case, and the game was fair.]

14. Let $a = 8$, $b = 4$, as in the text, and solve the game. Compare your solution with that in the text.

[*Ans.* Each player plays more conservatively; game is slightly more favorable to R than in the previous version.]

15. The players have agreed that the ante will be \$4. They are debating the size of the raise. What value of b should player R argue for? (*Hint:* He does not want the game to be fair. Then what are the possible values of b? Find the value of the 2×2 game for any such b, and find its maximum value by trying several values of b.)

SUGGESTED READING

von Neumann, J., and O. Morgenstern, *Theory of Games and Economic Behavior*, Princeton University Press, Princeton, 1944, 3d ed., 1953, Chap. I.

Williams, J. D., *The Compleat Strategyst*, McGraw-Hill, New York, 1950.

Kuhn, H. W., and A. W. Tucker, The article on the "Theory of Games," *Encyclopaedia Britannica*, 1956 edition.

Morgenstern, O., "The Theory of Games," *Scientific American*, **180** (1950), pp. 294-308.

McKinsey, J. C. C., *Introduction to the Theory of Games*, McGraw-Hill, New York, 1952.

Bellman, R., and D. Blackwell, "Red Dog, Blackjack, Poker," *Scientific American*, **184** (1951), pp. 44-47.

MacDonald, J., *Strategy in Poker, Business and War*, Norton, New York, 1950.

Charnes, A., and W. W. Cooper, "Linear Programming," *Scientific American*, **191** (1954), pp. 21-23.

Chapter VII*

APPLICATIONS TO BEHAVIORAL SCIENCE PROBLEMS

1. SOCIOMETRIC MATRICES

Matrices having only the entries 0 and 1 have been used by sociologists to analyze the structure of dominance relations in groups of subjects (animal or human). We shall use the notation $A_1 \gg A_2$ to indicate that individual A_1 "dominates" individual A_2. For example, in the pecking order of chickens in a barnyard, A_1 dominates A_2, that is, $A_1 \gg A_2$ means "chicken A_1 *pecks* chicken A_2." As another example, suppose that A_1 and A_2 are athletic teams and the relation $A_1 \gg A_2$ means "team A_1 *beats* team A_2."

We shall say that the relation $\gg$ is a *dominance relation* if it satisfies the following two properties:

(i) It is false that $A_i \gg A_i$; that is, no individual can dominate himself.

(ii) For each pair of individuals A_1 and A_2 either $A_1 \gg A_2$ or $A_2 \gg A_1$, but not both; that is, in every pair of individuals, there is exactly one who is dominant.

It has been observed that in the pecking order of chickens a dominance relation holds. Also, in the play of one round of a round robin contest among athletic teams, if ties are not allowed (as in baseball), then a dominance relation holds.

The reader may have been surprised that we did not assume that, if $A_1 \gg A_2$ and $A_2 \gg A_3$, then $A_1 \gg A_3$. This is the so-called transitive

law. A moment's reflection shows that the transitive law need not hold for dominance relations. Thus if team A beats team B and team B beats team C (in football, say), then we cannot assume that team A will necessarily beat team C. In almost every football season there are examples where "upsets" occur.

A convenient way of depicting dominance relations is by means of directed graphs. Two such directed graphs are shown in Figure 1.

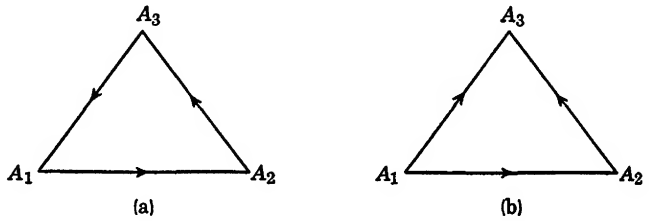


Figure 1

Individuals are represented on the graph as (lettered) points and a dominance relation between two individuals as a directed line segment (line segment with an arrow) connecting the two individuals. The graph in Figure 1(a) represents the situation: A_1 dominates A_2, also A_2 dominates A_3, and A_3 dominates A_1. Similarly, the graph in Figure 1(b) represents the situation: A_1 dominates A_2 and A_3, and A_2 dominates A_3. These graphs represent the two essentially different dominance relationships that are possible among three individuals (cf. Exercise 1).

Still another way in which dominance relations can be exhibited is by means of matrices, called *dominance matrices*, having only zeros and ones as entries. Two such matrices are shown in Figure 2.

$$\begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \end{array}\begin{array}{c} \begin{array}{ccc} A_1 & A_2 & A_3 \end{array} \\ \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \end{array} \qquad \begin{array}{c} \\ A_1 \\ A_2 \\ A_3 \end{array}\begin{array}{c} \begin{array}{ccc} A_1 & A_2 & A_3 \end{array} \\ \begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix} \end{array}$$

(a) (b)

Figure 2

Notice that we have labeled both the rows and the columns with letters of the individuals. An entry of 1 in the row of individual A_i and the column of individual A_j means that individual A_i dominates

A_j; that is, $A_i \gg A_j$. Similarly, a 0 entry there means that A_i does not dominate A_j. The reader can check that the dominance situations in Figures 2(a) and (b) are the same as those in Figures 1(a) and (b), respectively.

Since a dominance matrix is derived from a dominance relation, we can investigate the effects of conditions (i) and (ii) above on the entries in the matrix. Condition (i) simply means that all entries on the main diagonal (the one which slants downward to the right) of the matrix must be zero. Condition (ii) means that, whenever an entry above the main diagonal of the matrix is 1, the corresponding entry of the matrix which is placed symmetrically to it through the main diagonal is 0, and vice versa. To state these conditions more precisely, suppose that there are n individuals, and let D be a dominance matrix with entries d_{ij}. Then the conditions above are

(i) $d_{ii} = 0$ for $i = 1, 2, \ldots, n$.
(ii) If $i \neq j$, then $d_{ij} = 1$ if and only if $d_{ji} = 0$.

The 1 entries in the ith row correspond to the individuals whom A_i dominates, and the 1 entries in the jth column correspond to the individuals who dominate A_j.

Since a dominance matrix D is square, we can compute the powers of the matrix, D^2, D^3, etc. Let $E = D^2$, and consider the entry in the ith row and jth column of E. We have

$$e_{ij} = d_{i1}d_{1j} + d_{i2}d_{2j} + \ldots + d_{in}d_{nj}.$$

Now a term of the form $d_{ik}d_{kj}$ can be nonzero only if both factors are nonzero; that is, only if both factors are equal to 1. But if $d_{ik} = 1$, then individual A_i dominates A_k; and if $d_{kj} = 1$, then individual A_k dominates A_j. In other words, $A_i \gg A_k \gg A_j$. We shall call a dominance of this kind a *two-stage dominance*. (To keep ideas straight, let us call $A_i \gg A_j$ a *one-stage dominance*.) We now can see that the entry e_{ij} gives the number of two-stage dominances that individual A_i has over individual A_j.

For example, let D be the matrix

$$D = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Then D^2 is the matrix

$$D^2 = \begin{pmatrix} 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Thus we see that in this example A_1 has one two-stage dominance over A_3 and two two-stage dominances over A_4; similarly, A_2 has one two-stage dominance over A_4. These can be written down explicitly as

$$A_1 \gg A_2 \gg A_3,$$
$$A_1 \gg A_2 \gg A_4,$$
$$A_1 \gg A_3 \gg A_4,$$
$$A_2 \gg A_3 \gg A_4.$$

The directed graph for this dominance situation is given in Figure 3.

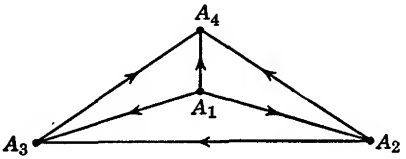


Figure 3

The reader should trace out on the graph of Figure 3 the two-stage dominances given above.

The following theorem will be proved in the next section.

Theorem. Let $\gg$ be a dominance relation on a set of n individuals $A_1, A_2, \ldots, A_n$. Then there exists at least one individual who can dominate in either one or two stages every other individual in the group. Also there exists at least one individual who is dominated in either one or two stages by every other individual in the group.

In matrix language the theorem can be restated as follows: let $S = D + D^2$; then there are at least one row and one column of S having all but one entry (the diagonal one) nonzero.

To illustrate this theorem consider the dominance situation shown in Figure 4. The dominance matrix D and its square D^2 for this situation are

$$D = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}, \qquad D^2 = \begin{pmatrix} 0 & 1 & 2 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

The matrix S corresponding to these is

$$S = D + D^2 = \begin{pmatrix} 0 & 2 & 2 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 1 & 1 & 2 & 0 \end{pmatrix}.$$

Observe that A_1, A_3, and A_4 can each dominate every other individual in one or two stages, but that A_2 cannot so dominate A_4. Similarly,

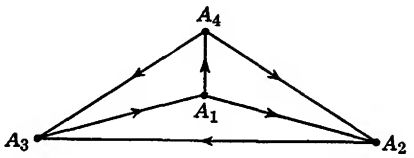


Figure 4

each of the individuals A_1, A_2, and A_3 is dominated in one or two stages by every other individual, while A_4 is not so dominated by A_2. It is instructive to check these statements in the directed graph of Figure 4.

As a final application of these dominance matrices, we shall define the power of an individual. By the *power* of an individual in a dominance situation, we mean the total number of one-stage and two-stage dominances which he can exert. Since the total number of one-stage dominances exerted by A_i is the sum of the entries in row i of the matrix D, and the total number of two-stage dominances exerted by A_i is the sum of the entries in row i of the matrix D^2, we see that the power of A_i can be expressed as:

The power of A_i is the sum of the entries in row i of the matrix $S = D + D^2$.

In the example of Figure 4 it is easy to check that the powers of the various individuals are the following:

The power of A_1 is 5.
The power of A_2 is 2.
The power of A_3 is 3.
The power of A_4 is 4.

Example. (Athletic contest). The idea of the power of an individual can be used to judge athletic events. For example, the result of a single round of a round robin athletic event results in the following data.

Team A beats teams B and D.
Team B beats team C.
Team C beats team A.
Team D beats teams C and B.

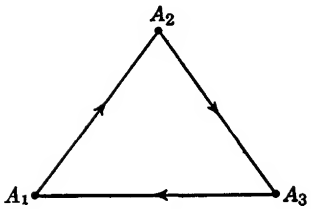
Then it is easy to check that this is precisely the dominance situation shown in Figure 4. By the analysis given above we can rate the teams in the following order according to their respective powers: A, D, C, and B.

It should be remarked that the above definition of the power of an individual is not the only one possible. In Exercise 10 below we suggest another definition of power which gives different results. Before using one or the other of these definitions, a sociologist should examine them carefully to see which (if either) fits his needs.

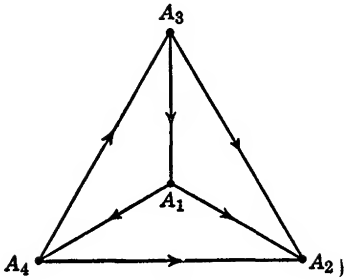
EXERCISES

1. Show that there are only two essentially different pecking orders possible among three chickens, namely, those given in Figure 1. (*Hint:* Use directed graphs.)

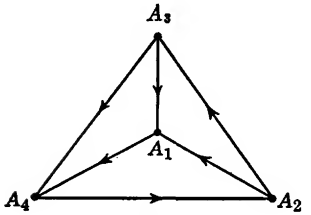
2. Find the dominance matrices D corresponding to the following directed graphs.



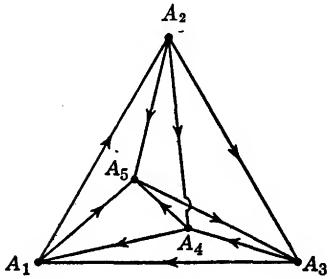
(a)



(b)



(c)



(d)

[*Ans.* (b) $\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix}$.]

3. Compute the matrices D^2 and $S = D + D^2$ and determine the powers of each of the individuals in the examples of Exercise 2.

[*Ans.* (b) $D^2 = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}$; $S = \begin{pmatrix} 0 & 2 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 2 & 1 & 0 \end{pmatrix}$; 4, 0, 4, 4.]

4. Find the powers of each of the individuals in the dominance situation whose matrix is

$$D = \begin{pmatrix} 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

[*Ans.* The powers are: 14, 14, 14, 14, 4, 10, 6, for A_1 through A_7, respectively.]

5. Find all the essentially different pecking orders that are possible among four chickens. [*Ans.* There are four essentially different ones.]

6. If D is any dominance matrix, give the interpretation of the entries in the columns of the matrix $S = D + D^2$. Also give the interpretation for the column sums of S.

7. If D is any dominance matrix, give the interpretation for the entries in the matrix D^3; also give the interpretations for the row and column sums of D^3; do the same for the entries and the row and column sums of the matrix $S = D + D^2 + D^3$.

[*Ans.* The entries in the ith row of D^3 give the three-stage dominances that A_i exerts; the ith row sum of D^3 gives the total number of three-stage dominances that A_i exerts. The entries in S give the one, two, or three-stage dominances, and the ith row sum of S gives the total number of such that A_i exerts.]

8. If D is any dominance matrix, give the interpretation for the entries in the matrix

$$S = D + D^2 + D^3 + \ldots + D^n.$$

Also give the interpretation for the row and column sums of this matrix.

9. A round robin tennis match among four people has produced the following results:

Smith has beaten Brown and Jones.
Jones has beaten Brown.
Taylor has beaten Smith, Brown, and Jones.

By finding the powers of each player, rank them into first, second, third, and fourth place. Does this ranking agree with your intuition?

[*Ans.* Taylor has power = 6, Smith has power = 3, Jones has power = 1, and Brown has power = 0.]

10. Let the power_1 of an individual be the power as defined in the text above. Define a new power, called power_2, of an individual as follows: If D is the dominance matrix for a group of n individuals, then the power_2 of A_i is the sum of row i of the matrix

$$S' = D + \tfrac{1}{2}D^2.$$

Find the power_2 of each of the teams in the athletic team example in the text. Show that the power_2 of a team need not equal his power_1. Comment on the result.

11. Find the power_2 of the players in Exercise 9. Discuss its relation with the power_1 of each of the players.

[*Ans.* Taylor has $\text{power}_2 = \frac{7}{2}$, Smith has $\text{power}_2 = 3$, Jones has $\text{power}_2 = 2$, Brown has $\text{power}_2 = \frac{3}{2}$.]

12. Discuss and give interpretations for the entries of the matrix

$$S = D + \tfrac{1}{2}D^2 + \tfrac{1}{3}D^3 + \ldots + (1/m)D^m.$$

Give interpretations for the row and column sums of this matrix.

13. Use the result of Exercise 5 to show that if a round robin tournament of four players is judged by their power, one of two things must happen: Either there is no tie, or there is a three-way tie. In the case of three-way ties show, by symmetry, that no rational criterion can be used to break these ties (without having playoff games, of course).

14. In the example in the text replace "beats" by "is beaten by." Does this reverse the order of the teams according to their power? Do the same for Exercise 9. [*Ans.* No; *CBAD;* Yes.]

2. COMMUNICATION NETWORKS

A communication network consists of a set of n people, call them $A_1, A_2, \ldots, A_n$, such that between some pairs of persons there is a communication link. A communication link may be either one-way or two-way. A two-way communication link might be made by telephone or radio, and a one-way link by sending a messenger, lighting a signal light, setting off an explosion, etc. We shall again use the symbol $\gg$ where $A_i \gg A_j$ now shall mean that individual A_i can communicate with A_j (in that direction). The only requirement that we now put on the symbol $\gg$ is the following:

(i) It is false that $A_i \gg A_i$ for any i; that is, an individual cannot (or need not) communicate with himself.

Notice that we have dropped the second condition which we used in the preceding section—we do not require that, of every pair of individuals, at least one can communicate with the other. Also, it is possible that both $A_i \gg A_j$ and $A_j \gg A_i$; that is, a two-way communication link is possible.

Again it is convenient to use directed graphs to represent communication networks. In Figure 5 we have drawn two such. The arrows, of course, indicate the direction in which communication is possible. A double arrow indicates communication is possible in both directions.

As before, we can also represent communication networks by means of matrices C having only 0 and 1 entries, which we call *communica-*

tion matrices. The entry in the ith row and jth column of C will be equal to 1 if A_i can communicate with A_j (in that direction) and

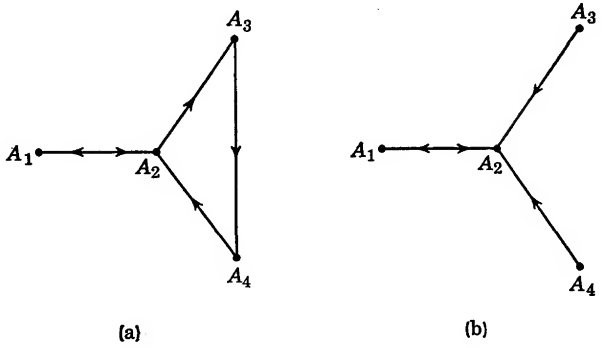


Figure 5

otherwise equal to 0. Thus the communication matrices corresponding to the communication networks of Figure 5 are given in Figure 6.

$$C = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad C = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

(a) (b)

Figure 6

Notice that the diagonal entries of the matrices in Figure 6 are all equal to 0. This is true in general for a communication matrix, since a restatement of condition (i) in matrix language is that $d_{ii} = 0$ for all i. It is not hard to see that any matrix having only 0 and 1 entries, and with all zeros down the main diagonal, is the communication matrix of some network.

The square of a communication matrix has an interpretation similar to the interpretation of the square of a dominance matrix. If C is a communication matrix, the entry in the ith row and jth column of C^2 gives the number of two-stage communications between A_i and A_j. For example, the square of the matrix in Figure 6(a) is

$$C^2 = \begin{pmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}.$$

The entry of 1 in the upper left-hand corner indicates that A_1 can communicate "with himself" in two stages. This is indeed true, the communication chain being $A_1 \gg A_2 \gg A_1$. One can also see, for example, that A_4 has two-stage communications with both A_1 and A_3. These and the other two-stage communications indicated in C^2 can easily be seen on the graph of Figure 5(a).

As in Section 1 we shall be interested in the sum of the matrices C and C^2. Let $S = C + C^2$. The following theorem includes the theorem in Section 1 (see Exercise 7).

Theorem. Let a communication network of n individuals be such that, for every pair of individuals, at least one can communicate in one stage with the other. Then there is at least one person who can communicate with every other person in either one or two stages. Similarly, there is at least one person who can be communicated with in one or two stages by every other person.

Stated in matrix language, the above theorem is: Let C be the communication matrix for the network described above; then there is at least one row of $S = C + C^2$ which has all its elements nonzero, except possibly the entry on the main diagonal. Similarly, there is at least one column having this property.

Proof. We shall prove only the first statement since the proof of the second is analogous.

First we shall prove the following statement: If A_1 cannot communicate in either one or two stages with A_i, where $i \neq 1$, then A_i can communicate in one stage with at least one more person than can A_1. We prove this in two steps. First by the hypothesis of the theorem we see that:

(a) If it is false that $A_1 \gg A_i$, then $A_i \gg A_1$. Second we can prove that:

(b) Suppose that for all k it is false that $A_1 \gg A_k \gg A_i$; it follows that, if $A_1 \gg A_k$, then also $A_i \gg A_k$.

For if $A_1 \gg A_k$, it is false that $A_k \gg A_i$; hence, by the hypothesis of the theorem, it is true that $A_i \gg A_k$.

Now (b) says that every one-stage communication possible for A_1 is also possible for A_i. From this and (a), it then follows that A_i can make at least one more (one-stage) communication than can A_1.

We now return to the proof of the theorem. Let $r_1, r_2, \ldots, r_n$ be the row sums of the matrix C. By renaming the individuals, if necessary, we can assume that the largest row sum is r_1, that is, $r_1 \geq r_k$ for $k = 1, 2, \ldots, n$. We shall show that A_1 can communicate with everyone else in one or two stages. (The proof is based on the indirect method.) Suppose, on the contrary, that there is an individual A_i, where $i > 1$, with whom A_1 cannot so communicate. By the statement proved above, A_i can communicate in one stage with at least one more person than A_1 can. But this implies that $r_i > r_1$, which contradicts the fact that we have named the individuals so that $r_1 \geq r_i$. This contradiction establishes the theorem.

An additional conclusion which can be made from the proof of the theorem is that the individual or individuals having the *largest* row sum in the matrix C can communicate with everyone else in one or two stages. Similarly, the individuals having the *largest* column sum can be communicated with by everyone in one or two stages.

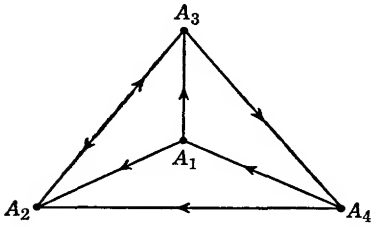


Figure 7

The network shown in Figure 7 satisfies the hypothesis of the theorem, hence its conclusion. The communication matrix for this network is

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 0 \end{pmatrix}.$$

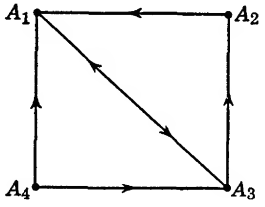
Here the maximum row sum of 2 occurs in rows one, three, and four, so that A_1, A_3, and A_4 can communicate with everyone else in one or two stages. (Find the necessary communication paths in Figure 7.) However, it requires three stages for A_2 to communicate with A_1. The maximum sum of 3 occurs in column two so that A_2 can be communicated with by everyone else in one or two stages (actually one stage is enough). It happens also that A_3 and A_4 can also be com-

municated with in one or two stages; however, as observed above, A_1 cannot be.

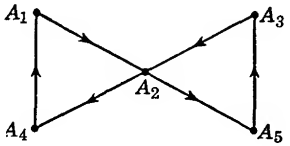
Neither of the networks in Figure 5 satisfies the hypothesis of the theorem. It happens that the network in Figure 5(a) does satisfy the conclusion of the theorem, while the network in Figure 5(b) does not. (See Exercise 4.)

EXERCISES

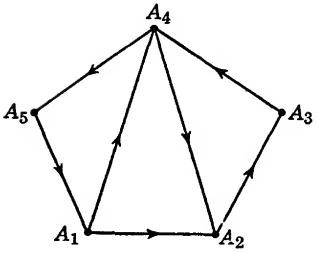
1. Find the communication matrices for the following communication networks.



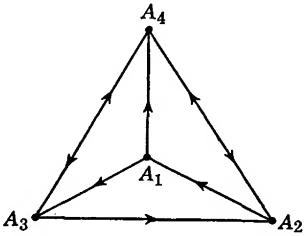
(a)



(b)



(c)



(d)

[*Ans.* (a) $\begin{pmatrix} 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 \end{pmatrix}$.]

2. Draw the directed graphs corresponding to the following communication matrices.

(a) $\begin{pmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$, (b) $\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$,

(Cont. on page 320)

(c) $\begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \end{pmatrix}$, (d) $\begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}$.

3. Which of the communication networks whose matrices are given in Exercise 2 satisfy the hypothesis of the theorem of this section?

[*Ans.* (a) and (c).]

4. Show that the network in Figure 5(a) satisfies the conclusion of the theorem, while the network in Figure 5(b) does not.

5. By computing the matrix S in each case, find the persons who can communicate with everyone else in one or two stages and those who can be communicated with in one or two stages, for the communication matrices in Exercise 2. (In some cases such persons need not exist. See Exercise 3.)

[*Ans.* (a) Everyone. (b) Everyone. (d) Neither type of person exists.]

6. Find all communication networks among three individuals which satisfy the hypothesis of the theorem of this section. How many of these are essentially different? [*Ans.* There are seven.]

7. Show that the theorem stated in the last section follows from the theorem proved in the present section.

8. If C is a communication matrix, give an interpretation for the entries of the matrix C^3. Do the same for the matrix C^4.

[*Ans.* The entry in row i and column j of C^3 gives the number of three-stage communications from i to j; the same entry of C^4 gives the number of four-stage communications from i to j.]

9. If C is a communication matrix, give an interpretation for the entries of the matrix $S = C + C^2 + C^3 + \ldots + C^m$.

10. Prove the second statement of the theorem of the present section.

11. Prove that the following statement is true: In a communication network involving three individuals, it is possible for a message starting from any person to get to any other person if and only if the following condition is satisfied: each individual can send a message to at least one person and can receive a message from at least one person.

12. Show that the matrix form of the condition in Exercise 11 is: every row and column of the communication matrix must have at least one nonzero entry.

13. Is the statement in Exercise 11 true for a communication network involving two individuals? For four or more individuals? [*Ans.* Yes; no.]

3. STOCHASTIC PROCESSES IN GENETICS

The simplest type of inheritance of traits in animals occurs when a trait is governed by a pair of genes, each of which may be of two types, say G and g. An individual may have a GG combination or Gg (which is genetically the same as gG) or gg. Very often the GG and Gg types are indistinguishable in appearance, and then we say that the G gene *dominates* the g gene. An individual is called *dominant* if he has GG genes, *recessive* if he has gg, and *hybrid* with a Gg mixture.

In the mating of two animals, the offspring inherits one gene of the pair from each parent, and the basic assumption of genetics is that these genes are selected at random, independently of each other. This assumption determines the probability of every type of offspring. Thus the offspring of two dominant parents must be dominant, of two recessive parents must be recessive, and of one dominant and one recessive parent must be hybrid. In the mating of a dominant and a hybrid animal, the offspring must get a G gene from the former and has probability $\frac{1}{2}$ for getting G or g from the latter, hence the probabilities are even for getting a dominant or a hybrid offspring. Again in the mating of a recessive and a hybrid, there is an even chance of getting either a recessive or a hybrid. In the mating of two hybrids, the offspring has probability $\frac{1}{2}$ for getting a G or a g from each parent. Hence the probabilities are $\frac{1}{4}$ for GG, $\frac{1}{2}$ for Gg, and $\frac{1}{4}$ for gg.

Example 1. Let us consider a process of continued crossings. We start with an individual of unknown genetic character, and cross it with a hybrid. The offspring is again crossed with a hybrid, etc. The resulting process is a Markov chain. The states are "dominant," "hybrid," and "recessive." The transition probabilities are

$$P = \begin{array}{c} \\ d \\ h \\ r \end{array} \begin{array}{c} \begin{array}{ccc} d & h & r \end{array} \\ \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & \frac{1}{2} & \frac{1}{2} \end{pmatrix} \end{array} \tag{1}$$

as can be seen from the previous paragraph. The matrix P^2 has all entries positive (see Exercise 1), hence we know from Chapter V, Section 7, that there is a unique fixed point probability vector, i.e., a vector p such that $pP = p$. By solving three equations, we find the fixed vector to be $p = (\frac{1}{4}, \frac{1}{2}, \frac{1}{4})$. Hence, no matter what type the origi-

nal animal was, after repeated crossing we have probability nearly $\frac{1}{4}$ of having a dominant, $\frac{1}{2}$ of having a hybrid, and $\frac{1}{4}$ of having a recessive offspring.

Example 2. If we keep crossing the offspring with a dominant animal, the result is quite different. The transition probabilities are

$$P' = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 1 & 0 \end{pmatrix}. \tag{2}$$

No power of this is all positive, and hence our general theorem does not apply. If we solve for p, we find that (1,0,0) is a unique probability vector fixed point. But here the components are not all positive. Therefore, after sufficiently long time has elapsed, we can be almost certain that we have a dominant offspring. This is easy to verify. Even if we start with a recessive animal, after a single crossing, the offspring cannot be recessive. It may be hybrid, but the probability of having a hybrid n times in a row is $(\frac{1}{2})^n$, which tends to zero. And once we have a dominant offspring, all future animals will be dominant. The analysis for crossing with recessive animals is very similar (see Exercise 2).

We can interpret our results for crossings of large numbers of animals. If a given population is crossed with hybrids, and the offspring are all crossed with hybrids, etc., then eventually we will have approximately $\frac{1}{4}$ dominants, $\frac{1}{2}$ hybrids, and $\frac{1}{4}$ recessives. While if we keep crossing them with dominants, then after sufficiently many crossings we can expect only dominants.

In Example 1 we may ask a more difficult question. Suppose that we have a regular matrix P (as in Example 1), with states $s_1, \ldots, s_n$. The process keeps going through all the states. If we are in s_i, how long, on the average, will it take for the process to return to s_i? We can even ask the more general question of how long, on the average, it takes to go from s_i to s_j.

The average here is taken in the sense of an expected value. There is a probability p_1 that we reach s_j for the first time in one step, p_2 that we reach it first in two steps, etc. The expected value is $p_1 \cdot 1 + p_2 \cdot 2 + \ldots$. (See Chapter IV, Section 12.) This, in general, requires a difficult computation. However, there is a much simpler

way of finding the expected values. Let the expected number of steps required to go from stage s_i to s_j be m_{ij}. How can we go from s_i to s_j? We go from s_i to s_k with probability p_{ik} which is one step. If $k = j$, we are there. If $k \neq j$, it takes an average of m_{kj} steps more. Hence m_{ij} is the sum of $p_{ik}m_{kj}$ for all $k \neq j$, plus 1. Let $\overline{M}$ be the matrix obtained from M by placing zeros in all the diagonal entries m_{ii}, and let C be the square matrix having all entries equal to 1. Then our equations can be replaced by the single matrix equation

$$M = P\overline{M} + C. \tag{3}$$

The m_{ij} can be found by solving these simultaneous equations (see Exercise 8). Let us multiply both sides of the equation by P^n. Observe that $PC = C$; hence $P^nC = C$ (see Exercise 9).

$$P^nM = P^{n+1}\overline{M} + C. \tag{4}$$

We know that, for a regular P, P^n approaches a matrix each of whose rows is the fixed point vector p. Hence all the rows of (4) approach the same vector equation,

$$pM = p\overline{M} + (1, \ldots, 1) \tag{5}$$

or

$$p(M - \overline{M}) = (1, \ldots, 1). \tag{6}$$

But all components of $M - \overline{M}$ except the diagonal ones are 0. Hence our equation simply states that $p_i m_{ii} = 1$ for each i. This tells us that $m_{ii} = 1/p_i$. *The average time it takes to return from s_i to s_i is the reciprocal of limiting probability of being in s_i.* In Example 1 this means that if we have a dominant offspring we will have another dominant in an average of four steps, after a hybrid we have another hybrid in an average of two steps, and a recessive follows a recessive on the average in four steps.

Example 3. A more interesting, and also more complex, process is obtained by crossing a given population with itself, and then crossing the offspring with offspring, etc. Let us suppose that our population has a fraction d of dominants, h hybrids, and r recessives. Then $d + h + r = 1$. If the population is very large and they are mated at random, then (by the law of large numbers) we can expect d^2 to be the fraction of matings in which both parents are dominant, $2dh$

the fraction of mating a dominant with a hybrid, etc. Hence we have simple formulas for the (approximate) fraction of offspring of various types. We will compute the fraction of dominants as an example.

To have a dominant offspring we must have mated two dominant parents, or a dominant and a hybrid parent, or two hybrids. In the first case we always get a dominant offspring, in the second the probability is $\frac{1}{2}$, in the third case it is $\frac{1}{4}$. Hence the fraction of dominant offspring is

$$d^2 + \tfrac{1}{2}\cdot 2dh + \tfrac{1}{4}h^2 = d^2 + dh + \tfrac{1}{4}h^2.$$

If we represent the fractions in a given generation by a row vector, the process may be thought of as a transformation T which changes a row vector into another row vector.

$$(7) \quad (d,h,r)\cdot T = (d^2 + dh + \tfrac{1}{4}h^2,\ dh + rh + 2dr + \tfrac{1}{2}h^2,\ r^2 + rh + \tfrac{1}{4}h^2).$$

The trouble is that (see Exercise 3) the transformation T is not linear. Nevertheless, we know that after n crossings the distribution will be $(d,h,r)T^n$, so that, if we can get a simple formula for T^n, we can describe the results simply. And here luck is with us.

Let us compute T^2, i.e., find what happens if we apply twice the transformation specified above. The first generation of offspring is distributed according to the formula (7). We now take the first component on the right side as d, the second as h, and the third as r, and compute $d^2 + dh + \frac{1}{4}h^2$, etc. Here we find to our surprise that $T^2 = T$. Hence $T^n = T$.

This means that $(d,h,r)T = (d,h,r)T^n$, which in turn means that the distribution after many generations is the same as in the first generation of offspring. Hence we say that the process reaches an *equilibrium* in one step. It must, however, be remembered that our fractions are only approximate, and are a good approximation only for very large populations.

For the geneticist, this result is very interesting. It shows that, in a population in which no mutations occur and selection does not take place, "evolution" is all over in a single generation.

To the mathematician the process is interesting since it is an example of a quadratic transformation, a transformation more complex than the linear ones we have heretofore studied.

EXERCISES

1. From (1) compute P^2, P^3, P^4, and P^5. Verify that $P^2 > 0$ and that the powers approach the expected form (see Chapter V, Section 7)

2. Set up the matrix corresponding to P' in (2) for the case of repeated crossing with recessive animals. Find the fixed point probability vector, and interpret it.

3. Prove that T is not a linear transformation. (*Hint:* check the conditions on linearity given in Chapter V, Section 9, and show by means of an example that T does not have one of these properties.)

4. In the text we computed the first component of (7). Verify that the other two are correctly given.

5. Compute T^2 by taking the first component of (7) as d, the second as h, the third as r, and substituting into the formula (7). Making use of the fact that $d + h + r = 1$, show that $T^2 = T$.

6. A fixed point of T is a vector such that $(d,h,r)T = (d,h,r)$. Write the conditions that such a vector must satisfy, and give three examples of such fixed vectors. What is the genetic meaning of such a distribution?

[*Ans.* For example, $(\frac{1}{3}, \frac{4}{9}, \frac{4}{9})$.]

7. In the matrix P the second row is equal to the fixed point vector. What significance does this have?

8. For Example 1 write the matrix M with unknown entries m_{ij}. Write M by replacing m_{11}, $\overline{m}_{22}$, and m_{33} by zeros. Then solve the nine simultaneous equations given by (3), to find the m_{ij}. Check that $m_{ii} = 1/p_i$.

[*Ans.* $m_{11} = 4$; $m_{12} = 2$, $m_{13} = 8$.]

9. From the definition of a stochastic matrix (Chapter V, Section 7), prove that $PC = C$.

10. Prove that, if P is a regular $n \times n$ stochastic matrix having column sums equal to 1, then it takes an average of n steps to return from any state to itself. (Cf. Chapter V, Section 7, Exercise 8.)

11. It is raining in the Land of Oz. In how many days can the Wizard of Oz expect to go on a picnic? (Cf. Chapter V, Section 8, Exercise 1.) [*Ans.* 4.]

The remaining exercises develop a simpler method of treating the nonlinear transformation T, in the text above.

12. Let p be the ratio of G genes in the population, and $q = 1 - p$ the ratio of g genes. Express p and q in terms of d, h, and r.

[*Ans.* $p = d + \frac{1}{2}h$, $q = r + \frac{1}{2}h$.]

13. Suppose that we take all the genes in the population, mix them thoroughly, and select a pair at random for each offspring. Show, using the result of Exercise 12, that the resulting distribution of dominant, hybrid, and recessive individuals is precisely that given in (7).

[*Ans.* $(d,h,r)\cdot T = (p^2,2pq,q^2)$.]

14. If we write $(d,h,r)\cdot T = (d',h',r')$, show, using the result of Exercise 13, that $h'^2 = 4d'r'$.

15. Show that for equilibrium it is necessary that $h^2 = 4dr$.

16. Show that if $h^2 = 4dr$, then $p^2 = d$, $q^2 = r$, and $2pq = h$. Hence show that this condition is also sufficient for equilibrium.

17. Use the results of Exercises 14-16 to show that the population reaches equilibrium in one generation.

4. ABSORBING MARKOV CHAINS AND GENETICS

There was an essential difference between the results of the first two examples in the last section. In Example 1 the process could have been in any one of the three states after a long time, and all we knew was what the three probabilities were. These we were able to obtain from the fact that we had a unique probability vector fixed point with all its components positive. In that process the fixed point furnished all the interesting information.

In Example 2 the fixed point told us only that eventually the process would end up in the first state, and would stay there. The principal characteristic of the first state is that, once the process enters this state, it cannot leave it. Such a state is described as *absorbing*.

DEFINITION. A *state* in a Markov chain is an *absorbing state* if it is impossible to leave it.

DEFINITION. A *Markov chain* is called *absorbing* if (1) it has at least one absorbing state, and (2) from every state it is possible to go to an absorbing state (not necessarily in one step).

Theorem. In an absorbing Markov chain it is certain that the process will end up in one of the absorbing states.

We shall indicate only the basic idea of the proof of the theorem. From each nonabsorbing state, s_j, it is possible to reach an absorbing state. Let n_j be the minimum number of steps required to reach an

absorbing state, starting from state s_j. Let p_j be the probability that, starting from state s_j, the process will not reach an absorbing state in n_j steps. Then $p_j < 1$. Let n be the largest of the n_j and p be the largest of the p_j. The probability of not being absorbed in n steps is less than p, in $2n$ steps is less than p^2, etc. Since $p < 1$, these probabilities tend to zero.

For an absorbing Markov chain we consider three interesting questions: (a) What is the probability that the process will end up in a given absorbing state? (b) On the average, how long will it take for the process to reach an absorbing state? (c) On the average, how many times will the process be in each nonabsorbing state? The answer to all these questions depends, in general, on which state the process starts from.

If there are at least two absorbing states, s_1 and s_2, the answer to question (a) must depend on the starting position. If the process starts in s_1, it ends up in s_1 with probability 1, while, if the process starts in s_2, it ends up in s_1 with probability 0. On the other hand, if s_1 is absorbing and s_2 is not, the answers to questions (b) and (c) depend on the starting point. If the process starts in s_1, the answer to both questions is 0. But if it starts in s_2 it takes at least one step to reach an absorbing state, and it will be in at least one nonabsorbing state once (namely s_2). (See Exercise 1.)

Let P be the matrix of transition probabilities for an absorbing Markov chain. Let s be an absorbing state. We want to find the probability that the process ends up in s, and this depends on where it starts. Let d_i be the probability of the process ending up at s, if it starts at s_i. Form the column vector d having d_i as its ith component. From state i the process can go to state j with probability p_{ij}. If it does, there is a probability d_j of going on from state j to s. Hence $p_{ij}d_j$ is the probability of going from s_i to s via s_j. The sum of these terms gives the total probability of going from s_i to s. Hence we will have,

$$d_i = p_{i1}d_1 + p_{i2}d_2 + \ldots + p_{in}d_n,$$

for $i = 1, 2, \ldots, n$. But the sum on the right side is simply the ith component of Pd; hence we can shorten our equations to the single vector equation

$$d = Pd. \tag{1}$$

We note that d would be a "fixed point" of P, except that it is a column vector instead of a row vector. For an absorbing Markov chain there frequently is more than one probability vector solution of (1). We can select the one we want by imposing the conditions

(2) The component of d corresponding to s is 1; corresponding to any other absorbing state is 0.

This condition gives us a unique fixed vector d for every absorbing state, the vectors differing only with respect to (2).

Next we will study question (c), since the answer to question (b) will be a direct consequence. Let the process start in a nonabsorbing state s_i. How many times do we expect it to be in nonabsorbing state s_j? (Here "expect" is to be taken in the sense of Chapter IV, Section 12.) Let us call this number t_{ij}. From s_i, it moves to state s_k with probability p_{ik}. If s_k is an absorbing state, it will never come to s_j. If it is nonabsorbing, we expect it to come to s_j a total of t_{kj} times. Hence $p_{ik}t_{kj}$ has to be summed for all nonabsorbing states s_k. But we do not yet have the entire answer, for if $i = j$, then we must not forget that we started in s_j, and hence we must add 1. For s_i, s_j nonabsorbing,

$$t_{ij} = \underset{\substack{\text{summed over the} \\ \text{nonabsorbing states}}}{p_{i1}t_{1j} + \ldots + p_{in}t_{nj}} \quad (+1 \text{ if } i = j).$$

We can rewrite all these equations as a single matrix equation. Let us form the truncated matrix Q which is obtained from P by crossing out all rows and columns corresponding to absorbing states. Let a be the number of absorbing states; then Q is an $(n - a) \times (n - a)$ matrix. The matrix T, whose components are t_{ij}, which are defined only when s_i and s_j are nonabsorbing, is the same size as Q. The sum on the right side above is simply the product of row i of Q with column j of T; hence it is component ij of QT. To this we want to add 1 if $i = j$; hence we want to add the corresponding component of the identity matrix. This yields the matrix equation

$$T = QT + I. \tag{3}$$

We can rewrite equation (3) as

$$I = T - QT = (I - Q)T.$$

Multiplying both sides by $(I - Q)^{-1}$, we obtain the solution

$$T = (I - Q)^{-1}. \tag{4}$$

Thus we find that the components of $(I - Q)^{-1}$ provide the answer to question (c). They also provide the answer to question (b). Let us suppose that we want to know how many steps it takes the process on the average to reach an absorbing state from s_i (nonabsorbing). There will be one step for each time that it is in a nonabsorbing state. Hence the total number of times it is in a nonabsorbing state is the same as the number of steps it takes to reach an absorbing state. But this total is simply the number of times on the average that the process is in each nonabsorbing state, starting with s_i, which is the sum of row i of $(I - Q)^{-1}$. Hence the row sums of (4) provide the answer to question (b). Let us write this in vector form. Let t_i be the expected number of steps that take us from a nonabsorbing s_i to an absorbing state, and let t be the column vector having t_i as components. Then t consists of the row sums of (4). By using the column vector c having all components equal to one, we can write

$$t = (I - Q)^{-1}c. \tag{5}$$

Example 1. Let us return to equation (2) of the last section. P' represents a chain with one absorbing state, s_1. If we try to solve the equation $d = P'd$, we find that the solution can be any column vector all of whose components are the same. Using condition (2), we know that its first components must be 1; hence all components are 1. Thus

$$d = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

This means that no matter where the process is, the probability of ending up in s_1 is 1. This agrees with our first theorem, since there is only one absorbing state.

Next we form the truncated matrix

$$Q' = \begin{array}{c} \\ s_2 \\ s_3 \end{array} \begin{array}{c} \begin{array}{cc} s_2 & s_3 \end{array} \\ \begin{pmatrix} \frac{1}{2} & 0 \\ 1 & 0 \end{pmatrix} \end{array};$$

then $$(I - Q') = \begin{pmatrix} \frac{1}{2} & 0 \\ -1 & 1 \end{pmatrix};$$

and $$T = (I - Q')^{-1} = \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix}.$$

Hence we see that if the process starts in s_2, we can expect it to be in this state twice (including the starting position). If we are in s_3 we can expect it to be there only once, namely at the start, which is clear since it is impossible for it to return to s_3. And starting from s_3 we can expect it to be in s_2 twice.

$$t = \begin{pmatrix} 2 & 0 \\ 2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}.$$

Hence we expect it to reach s_1 in two steps from s_2 and in three steps from s_3.

Example 2. Let us construct a more complicated example of an absorbing Markov chain. We start with two animals of opposite sex, cross them, select two of their offspring of opposite sex and cross those, etc. To simplify the example we will assume that the trait under consideration is independent of sex.

Here a state is determined by a pair of animals. Hence the states of our process will be: $s_1 = (D,D)$, $s_2 = (D,H)$, $s_3 = (D,R)$, $s_4 = (H,H)$, $s_5 = (H,R)$, and $s_6 = (R,R)$. Let us illustrate the calculation of transition probabilities in terms of s_2. When the process is in this state, one parent has GG genes, the other Gg. Hence the probability of a dominant offspring or a hybrid offspring is $\frac{1}{2}$ for each. Then the probability of transition to s_1 (selection of two dominants) is $\frac{1}{4}$, transition to s_2 is $\frac{1}{2}$, and to s_4 is $\frac{1}{4}$. The transition matrix is

$$P'' = \begin{array}{c} \\ s_1 \\ s_2 \\ s_3 \\ s_4 \\ s_5 \\ s_6 \end{array} \begin{array}{c} \begin{array}{cccccc} s_1 & s_2 & s_3 & s_4 & s_5 & s_6 \end{array} \\ \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{4} & \frac{1}{2} & 0 & \frac{1}{4} & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ \frac{1}{16} & \frac{1}{4} & \frac{1}{8} & \frac{1}{4} & \frac{1}{4} & \frac{1}{16} \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \end{array}.$$

There are two absorbing states, s_1 and s_6. The probabilities of absorp-

tion in s_1 are found by solving $d = P''d$, with the conditions $d_1 = 1$ and $d_6 = 0$ giving

$$d = \begin{pmatrix} 1 \\ \frac{3}{4} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{4} \\ 0 \end{pmatrix}.$$

Since the sum of these vectors d must be c (see Exercise 6) it follows that the other vector, giving the probability of absorption in s_6 is $c - d$.

The genetic interpretation of absorption is that after a large number of inbreedings either the G or the g gene must disappear. It is also interesting to note in the vector d that the probability of ending up entirely with G genes, if we start from a given state, is equal to the proportion of G genes in this state.

Next we form the truncated matrix

$$Q'' = \begin{array}{c} \\ s_2 \\ s_3 \\ s_4 \\ s_5 \end{array} \begin{array}{c} \begin{array}{cccc} s_2 & s_3 & s_4 & s_5 \end{array} \\ \begin{pmatrix} \frac{1}{2} & 0 & \frac{1}{4} & 0 \\ 0 & 0 & 1 & 0 \\ \frac{1}{4} & \frac{1}{8} & \frac{1}{4} & \frac{1}{4} \\ 0 & 0 & \frac{1}{4} & \frac{1}{2} \end{pmatrix} \end{array},$$

and

$$I - Q'' = \begin{pmatrix} \frac{1}{2} & 0 & -\frac{1}{4} & 0 \\ 0 & 1 & -1 & 0 \\ -\frac{1}{4} & -\frac{1}{8} & \frac{3}{4} & -\frac{1}{4} \\ 0 & 0 & -\frac{1}{4} & \frac{1}{2} \end{pmatrix},$$

and

$$T = (I - Q'')^{-1} = \begin{pmatrix} \frac{8}{3} & \frac{1}{6} & \frac{4}{3} & \frac{2}{3} \\ \frac{4}{3} & \frac{4}{3} & \frac{8}{3} & \frac{4}{3} \\ \frac{4}{3} & \frac{1}{3} & \frac{8}{3} & \frac{4}{3} \\ \frac{2}{3} & \frac{1}{6} & \frac{4}{3} & \frac{8}{3} \end{pmatrix},$$

and

$$t = T\begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 4\frac{5}{6} \\ 6\frac{2}{3} \\ 5\frac{2}{3} \\ 4\frac{5}{6} \end{pmatrix}.$$

Hence we see that, if we start in a state other than (D,D) or (R,R), we can expect to reach one of these states in about five or six steps. The exact expected times are given by the entries of t. The matrix T provides more detailed information, namely how many times we can expect to have offspring of the types (D,H), (D,R), (H,H), and (H,R), starting from a given nonabsorbing state. And the vectors d and $c - d$ give the probabilities of ending up in s_1 or s_6, respectively. These quantities jointly give us an excellent description of what we can expect of our process.

EXERCISES

1. Prove that in an absorbing Markov chain:
   (a) The probability of reaching a given absorbing state is independent of the starting state if and only if there is only one absorbing state.
   (b) The expected time for reaching an absorbing state is independent of the starting state if and only if every state is absorbing.

2. Verify that the inverse $(I - Q')^{-1}$ is correctly given in the text.

3. Verify that the inverse $(I - Q'')^{-1}$ is correctly given in the text.

4. Solve the equation $d = P'd$ (see page 322).

5. Find two solutions of $d = P''d$, corresponding to absorption in s_1 and s_6, respectively. Verify that their sum is c (i.e., a vector all of whose components are 1).

6. Consider all vectors d which represent probabilities of absorption in a given absorbing state. Interpret the sum of two such vectors. Interpret the sum of all such vectors. What must the sum of all these vectors be?

7. Find two different probability vector fixed points of P''.
[*Ans.* (1,0,0,0,0,0); (0,0,0,0,0,1).]

8. There is an alternate method of computing T: we want to know the probability of being in s_j after n steps if we start in s_i, and sum this for all n. The sum of these will be t_{ij}.
   (a) Show that the probability after zero steps is given by I.
   (b) Show that the probability after $n > 0$ steps is Q^n.
   (c) Show that $T = I + Q + Q^2 + Q^3 + \ldots$.
   (d) Compute the sum of this series as if it were an ordinary geometric series.
   (e) Verify that the answer is the same as (4).

9. There is a simpler method of finding t. If t_i is the expected number of steps for reaching an absorbing state from s_i, this must be the same as taking one more step and then adding $p_{ij}t_j$ for every nonabsorbing state s_j.

(a) Give reasons for the above claim that $t_i = 1 +$ sum of $p_{ij}t_j$ over nonabsorbing states.

(b) Write these equations as a single vector equation.

(c) Solve for t.

(d) Verify that the solution is the same as (5).

10. Suppose that hybrids have a high mortality rate; say that half of the hybrids die before maturity, while only a negligible number of dominants and recessives die before maturity.

(a) In Example 2 above, modify the matrix P'' to apply to this situation.

(b) What are the absorbing states?

(c) Verify that it is an absorbing chain.

(d) Find the vectors d representing the probabilities of absorption in the various absorbing states.

[*Ans.* For s_1, $d = \begin{pmatrix} 1 \\ \frac{9}{10} \\ \frac{1}{2} \\ \frac{1}{2} \\ \frac{1}{10} \\ 0 \end{pmatrix}$.]

(e) Find T, and interpret.

(f) Find t, and interpret.

[*Ans.* $t = \begin{pmatrix} \frac{65}{26} \\ \frac{117}{26} \\ \frac{91}{26} \\ \frac{65}{26} \end{pmatrix}$.]

The remaining problems concern the inheritance of color-blindness, which is a sex-linked characteristic. There is a pair of genes, C and N, of which the former tends to produce color-blindness, the latter normal vision. The N gene is dominant. But a man has only one gene, and if this is C, he is color-blind. A man inherits one of his mother's two genes, while a woman inherits one gene from each parent. Thus a man may be of type C or N, while a woman may be of type CC or CN or NN. We will study a process of inbreeding similar to that of Example 2.

11. List the states of the chain. (*Hint:* There are six.)

12. Compute the transition probabilities.

13. Show that the chain is absorbing, and interpret the absorbing states. [*Ans.* In one the N gene disappears, in the other the C gene is lost.]

14. Prove that the probability of absorption in the state having only C genes, if we start in a given state, is equal to the proportion of C genes in that state.

15. Find T, and interpret.

16. Find t, and interpret.

[*Ans.* $\begin{pmatrix} 5 \\ 6 \\ 6 \\ 5 \end{pmatrix}$; if we start with both C and N genes, we can expect one of these to disappear in five or six crossings.]

5. THE ESTES LEARNING MODEL

In this section we shall discuss a mathematical model for learning proposed by W. K. Estes. We shall not give the most general theory, but only some special cases.

The theory was developed to explain certain kinds of learning which can be illustrated by experiments of the following kind. Suppose for example that a rat is put in a T maze (see Chapter III, Section 2) and goes either right or left. The experimenter places food on one side, and if the rat goes to the correct side he is rewarded. This experiment is then repeated many times, using some particular feeding schedule. The interest here lies in trying to predict the behavior of the rat under the different feeding schedules. For example, if the food is always placed on the right side, will the rat eventually learn this and always go right?

A similar experiment, performed with a human subject, is the following. A subject is given a sequence of heads and tails and each time is asked to guess what the next choice will be. He is to try to get as many right as possible. Again there are various ways that the experimenter can produce his sequences of H's and T's, and the interest lies in how the subject will react to different choices.

In the Estes model it is assumed that there are a finite number of elements, called "stimulus elements." At any given time each of these elements is connected either to a response R_0 or to a response R_1. These connections are allowed to change from experiment to experiment.

In a single experiment there is a certain probability θ that any particular stimulus element will be sampled by the subject. To say that an element is sampled is the same as to say that it has an effect

upon the subject on that experiment. It is assumed that elements sampled and connected to R_0 influence the subject in the direction of producing an R_0 response, and those sampled and connected to R_1 to produce an R_1 response.

The samplings of the various elements are assumed to be an independent trials process (see Chapter IV, Section 8). Thus, for example, if there are three stimulus elements a, b, and c, the probability that a is sampled, b is not sampled, and c is sampled would be $\theta(1 - \theta)\theta$.

We also assume that the experimenter takes one of two possible "reinforcing" actions, A_0 or A_1. This action may be taken before or after the subject's choice, but we assume that the subject learns of the choice of the experimenter only after he has made his own choice. In most experiments the subject would like to make R_0, if the experimenter makes A_0, and R_1 if the experimenter chooses A_1. We shall say that the subject "guesses correctly" if he matches the choice of the experimenter, i.e., does R_0 when the experimenter does A_0, or R_1 when the experimenter does A_1. In some experiments (e.g., the rat experiment above), he is rewarded if he does guess correctly the choice of the experimenter.

The following two basic assumptions are made:

Assumption A. The probability that the subject makes response R_1 is equal to the proportion of elements in the set sampled that are connected to R_1. If no elements are sampled, the responses are assumed equally likely.

Assumption B. If, in a given experiment, the experimenter chooses A_0, then all the elements that were sampled on this experiment, and that were connected to R_1, have their connections changed to R_0. If the experimenter chooses A_1, then all the elements sampled and connected to R_0 have their connections changed to R_1.

Note that in a single experiment only the set of elements that are actually sampled play a role, and these are the only elements whose connections can be changed by this experiment. In general, however, a different set will be sampled on each experiment, so that all the elements will at some time have an effect.

By assumptions A and B it is clear that the future choices of the subject are going to depend upon the choice of the experimenter. Therefore we must describe the method that the experimenter uses

to determine his A's. Typical schemes that have been used in actual experiments are the following:

(i) Choose A_0 with probability p, independent of the choice of the subject.

(ii) Make the same choice as the subject made (i.e., choose A_0 if he chose R_0, A_1 if he chose R_1).

(iii) Choose A_0 if the response of the subject on the previous experiment was R_0. Choose A_0 and A_1 with equal probabilities if his response was R_1.

We can describe a general class of schemes of the above kind as follows: We assume that the experimenter chooses A_1 with probability a, if the subject made response R_0 on the previous experiment; and chooses A_0 with probability b, if the subject made response R_1 on the last experiment. We can represent the choices of the experimenter for each choice of the subject by the matrix

$$\begin{array}{c} \\ R_0 \\ R_1 \end{array} \begin{array}{c} \begin{array}{cc} A_0 & A_1 \end{array} \\ \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \end{array}.$$

Thus in the above examples, (i) is the case $1 - a = b = p$, (ii) is

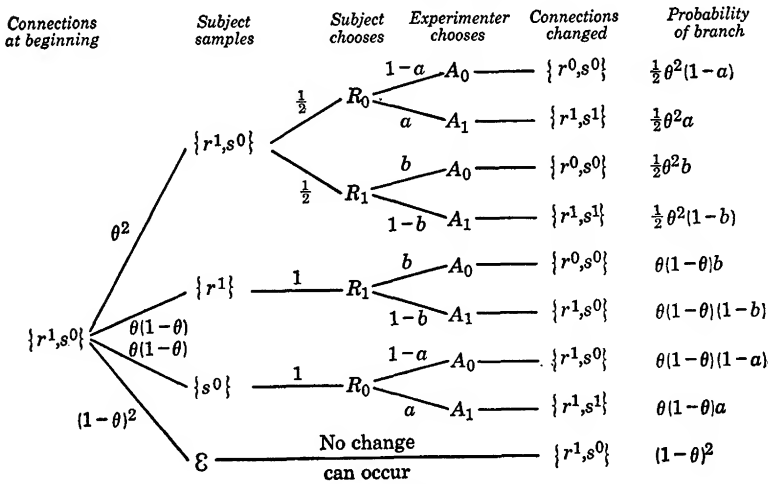


Figure 8

the case $a = 0$, $b = 0$, and (iii) is the case $a = 0$, $b = \frac{1}{2}$. We shall consider throughout this section and the next the case of two stimulus elements (called r and s). The analysis for a larger number of elements is similar but more complicated. Many of the results do not depend upon the number of stimulus elements assumed.

In Figure 8 we have indicated by a tree the various stages in a single experiment. We label the elements by superscripts 0 or 1 to indicate that an element is connected to R_0 or R_1. Let us assume that initially r is conditioned to R_1, and s to R_0; that is, we write $\{r^1,s^0\}$ as our starting point.

If we consider a sequence of experiments, we can obtain a Markov chain as follows. We take as states the number of elements connected to R_1 at any given time. Thus we can label the states 0, 1, and 2. Since all our probabilities depend only on the number of elements sampled, $\{r^1,s^0\}$ and $\{r^0,s^1\}$ may be thought of as the same state. This justifies us in taking the *number* of the elements conditioned to R_1 as representing our state. The transition probabilities are then found as follows. To find $p_{1,0}$ we look on our tree for all paths leading to connection $\{r^0,s^0\}$, and add their probabilities. Doing this, we obtain

$$p_{1,0} = \tfrac{1}{2}\theta^2(1 - a) + \tfrac{1}{2}\theta^2 b + \theta(1 - \theta)b.$$

We can also find from the tree $p_{1,1}$ and $p_{1,2}$. To find the other transition probabilities we would have to construct a similar tree, assuming that we started with a case where no elements were connected to R_1 and also a case where both elements were so connected. (See Exercise 1.) When this is done we obtain the complete matrix of transition probabilities.

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{pmatrix} (1-\theta)^2 a + 1 - a & 2\theta(1-\theta)a & \theta^2 a \\ \tfrac{1}{2}\theta^2(1-a) + \tfrac{1}{2}\theta(2-\theta)b & (1-\theta)^2 + \theta(1-\theta)(1-a) + \theta(1-\theta)(1-b) & \tfrac{1}{2}\theta^2(1-b) + \tfrac{1}{2}\theta(2-\theta)a \\ \theta^2 b & 2\theta(1-\theta)b & (1-\theta)^2 b + (1-b) \end{pmatrix} \end{array}.$$

In the next section we shall study this Markov chain in more detail.

EXERCISES

1. Construct a tree to show the possibilities for the connections after an experiment if the two stimulus elements are both connected to R_1 at the beginning of the experiment. Do the same for the case of no elements connected to R_1 at the beginning of the experiment.

2. Using the trees in Exercise 1, verify that the transition probabilities $p_{0,j}$ and $p_{2,j}$ given above are correct.

3. What is the probability that the subject will make response R_1 if at the beginning of the experiment one element is connected to each response? What is this probability if at the beginning of the experiment both elements are connected to response R_1? [*Ans.* $\frac{1}{2}$; $\frac{1}{2} + \theta - \frac{1}{2}\theta^2$.]

In the following exercises, find the matrix of transition probabilities under the special assumptions given in the problem. State whether the resulting Markov chain is absorbing or regular. Give an interpretation for each of the special cases in terms of the actual experiment. If the process is regular, find the limiting probabilities. If the process is absorbing, find the expected number of steps before absorption for each possible starting state. (See Section 4, Exercise 9.)

4. $a = 1$, $b = 1$, $\theta = \frac{1}{2}$. [*Ans.* Regular; (.3, .4, .3).]

5. $a = 1$, $b = 0$. [*Ans.* Absorbing; $t_0 = (3 - 2\theta)/(2\theta - \theta^2)$; $t_1 = 1/\theta$.]

6. $a = \frac{1}{4}$, $b = \frac{3}{4}$, $\theta = .1$.

7. $a = 0$, $b = \frac{1}{2}$, $\theta = \frac{1}{2}$.

8. $a = 1$, $b = \frac{1}{2}$, $\theta = \frac{1}{2}$.

9. $a = 0$, $b = 0$.

10. $\theta = 0$.

11. Consider the case $\theta = 1$, $0 < a < 1$, and $0 < b < 1$. Show that the matrix of transition probabilities is not regular and not absorbing.

12. In the case of Exercise 11, show that if the process starts in state 0 or 2, it never reaches state 1. Hence this state can be removed. Show that, if this is done, the resulting two-state Markov chain has a regular transition matrix. Find the limiting probability. [*Ans.* $\left(\frac{b}{a+b}, \frac{a}{a+b}\right)$.]

13. In the case of Exercise 11, show that there is a limiting probability of being in each of the states which is independent of the starting state. Find the limiting probabilities. Compare your answer with Exercise 12.

6. LIMITING PROBABILITIES IN THE ESTES MODEL

We wish now to study the limiting probabilities that the subject and that the experimenter will choose each of the possible alternatives. We are primarily interested in the effects of the stimulus elements. When the subject does not sample any stimulus element, they have no effect. Thus we shall calculate all our probabilities under the assumption that the subject samples at least one element. This amounts simply to throwing out all experiments on which the subject does not sample any element. Thus in the present section we shall assume that this has been done, and in any reference to experiments, it will be understood to mean an experiment on which the subject sampled at least one element.

With the above convention, if our process is in state 0 on a given experiment, then the probability that the subject will make response R_0 is (by assumption A) equal to 1. If it is in state 1, then by symmetry this probability is $\frac{1}{2}$. If it is in state 2, it is (by assumption A) equal to 0.

The matrix P will be regular if and only if the quantities a, b, θ, and $1 - \theta$ are all not zero (see Exercise 1). If the matrix is regular, then there will be a limiting probability for being in each of the states. These probabilities can be represented by a vector $p = (p_0,p_1,p_2)$ and found by solving the equations

$$pP = p.$$

If these equations are solved, we obtain

$$p_0 = \frac{b\theta + 2b^2(1 - \theta)}{(a + b)\theta + 2(a + b)^2(1 - \theta)},$$

$$p_1 = \frac{4ab(1 - \theta)}{(a + b)\theta + 2(a + b)^2(1 - \theta)},$$

$$p_2 = \frac{a\theta + 2a^2(1 - \theta)}{(a + b)\theta + 2(a + b)^2(1 - \theta)}.$$

From these probabilities we can find that the limiting probability that the subject will make response R_0 is

$$1 \cdot p_0 + \tfrac{1}{2}p_1 + 0 \cdot p_2 = \frac{b}{a + b},$$

and that the limiting probability that the subject makes response R_1 is $a/(a + b)$.

To find the probability that the experimenter makes the choice A_0, we must multiply the probabilities for each of the choices of the subject, by the probabilities that the experimenter does A_0 if the subject made the particular choice. Thus the limiting probability that the experimenter makes choice A_0 is

$$\frac{b(1 - a)}{a + b} + \frac{ab}{a + b} = \frac{b}{a + b}.$$

Thus we see that the limiting probability that the subject will make response R_0 is equal to the limiting probability that the experimenter will choose A_0. From these limiting probabilities we can find the limiting probability that the subject will guess correctly (see Exercise 3).

If we assume that the experimenter makes response A_0 with probability p independent of the choice of the subject, the subject can maximize the expected number of correct responses by always making response R_0 if $p > \frac{1}{2}$ and always making R_1 if $p < \frac{1}{2}$. (See Exercise 5.) The model predicts a less rational choice on the part of the subject. This would not seem disturbing in the case of the rat, but it would be hoped humans would do better. Unfortunately experiments have borne out that the model's predictions are approximately correct even with human subjects.

The following interesting experiment was performed by W. K.Estes and others with many types of subjects. If the subject does R_0, he is rewarded half the time; if he does R_1 he is never rewarded. One might expect that the subject will learn to do R_0, but this is not the case. What does the theory predict? If R_0 is chosen, reward follows half the time. Hence $a = \frac{1}{2}$. If R_1 is chosen, reward never follows. Hence $1 - b = 0$ or $b = 1$. The theory predicts a limiting probability of $b/(a + b) = \frac{2}{3}$ for the subject to choose R_0, which is in good agreement with experimental results.

We next consider an absorbing case. Specifically, we consider the case $a = 0$ and $b = 1$. This means that the experimenter always does A_0. The matrix of transition probabilities here is

$$P = \begin{pmatrix} 1 & 0 & 0 \\ \theta & 1 - \theta & 0 \\ \theta^2 & 2\theta(1 - \theta) & (1 - \theta)^2 \end{pmatrix}.$$

We shall use the methods developed in Section 4 to study this Markov chain. We have one absorbing state, namely, 0. Thus we know that the process will eventually enter this state and remain there. Being in this state means, by assumption A of the previous section, that the subject is sure to make response R_0. Thus being absorbed can be interpreted as the subject "learning" that the experimenter always does A_0.

We have seen that in an absorbing Markov chain it is possible to find the expected number of times that the process will be in each of the states before being absorbed, assuming some given starting state. Let t_{ij} be the expected number of times the process will be in state j if it starts in state i. Before calculating t_{ij} we consider what the knowledge of these quantities would tell us about the experiment. We observe that every time the process is in state 1, the subject chooses R_1 with probability $\frac{1}{2}$ and hence makes a wrong response with probability $\frac{1}{2}$. Every time the process is in state 2, the subject is sure to make response R_1, that is, to make a wrong response. Thus the expected number of wrong responses that the subject will make before learning is

$$\frac{1}{2}t_{i1} + t_{i2} \quad \text{for} \quad i = 1, 2, \tag{1}$$

assuming that the process starts in state i.

We find the t_{ij} as in Section 4. We first form the truncated matrix Q obtained from P by omitting the column and the row corresponding to the absorbing state.

$$Q = \begin{pmatrix} 1 - \theta & 0 \\ 2\theta(1 - \theta) & (1 - \theta)^2 \end{pmatrix}.$$

We then find $(I - Q)^{-1}$ to be

$$\begin{pmatrix} t_{11} & t_{12} \\ t_{21} & t_{22} \end{pmatrix} = (I - Q)^{-1} = \begin{pmatrix} \frac{1}{\theta} & 0 \\ \frac{2(1 - \theta)}{\theta(2 - \theta)} & \frac{1}{\theta(2 - \theta)} \end{pmatrix}.$$

Then from (1) we obtain $1/2\theta$ as the expected number of wrong responses if the process begins in state 1, and $1/\theta$ as the expected number of wrong responses if the process begins in state 2.

Of course it is true that in an actual experiment the starting state would not be known. However, it is not unreasonable to assume that on the first experiment the stimuli elements are connected at random.

This would mean that the process starts at state 0 with probability $\frac{1}{4}$, at state 1 with probability $\frac{1}{2}$, and at state 2 with probability $\frac{1}{4}$. Thus under this assumption the expected number of wrong responses before learning is

$$\frac{1}{2} \cdot \frac{1}{2\theta} + \frac{1}{4} \cdot \frac{1}{\theta} = \frac{1}{2\theta}. \tag{2}$$

EXERCISES

1. Prove that the matrix P in Section 5 is regular if and only if a, b, θ, and $1 - \theta$ are all different from zero. (*Hint:* Show that if any one of the quantities is 0, the chain is not regular.)

2. Verify that the probability that the subject makes response R_0 is $b/(a + b)$ by finding $1 \cdot p_0 + \frac{1}{2} \cdot p_1 + 0 \cdot p_2$.

3. Show that the limiting probability that the subject's choice agrees with that of the experimenter is

$$\frac{a(1 - b) + b(1 - a)}{a + b}.$$

4. Assume that the experimenter always chooses A_0 with a fixed probability p, independent of the choice of the subject. What proportion would the subject expect to get correct? [*Ans.* $1 - 2p + 2p^2$.]

5. Suppose under the conditions of Exercise 4 that the subject were always to make response R_0. Show that if $p > \frac{1}{2}$, then on the average the subject will do better by this method than by the method predicted by the model.

6. Consider the case $a = \frac{1}{2}$, $b = 0$, and $\theta = \frac{1}{2}$. For each possible starting state find the expected number of times that the process will be in each of the states before being absorbed. [*Ans.* $t_{00} = 3$; $t_{01} = 2$; $t_{10} = \frac{1}{2}$; $t_{11} = 3$.]

7. Do the same as in Exercise 6, for the case $a = 0$, and $b = 0$.

8. In Exercises 6 and 7 find the expected number of incorrect responses that the subject will make, assuming each possible starting state.

[*Ans.* 4,2,0; 0,0,0.]

9. In Exercises 6 and 7 find the expected number of incorrect responses that the subject will make assuming random connections for the stimuli elements on the first experiment, as in (2).

10. If the subject chooses R_0, he is rewarded with probability p. If he chooses R_1, he is never rewarded. (See the example with $p = \frac{1}{2}$ in the text above.) Find a and b. What is the limiting probability that the subject

chooses R_0? How often is he rewarded? How often would he be rewarded if he always chose R_0? Compare these two values for $p = \frac{3}{4}, \frac{1}{2}, \frac{1}{4}$.

[*Ans.* $1/(2-p)$; $p/(2-p)$; p.]

11. Compute p_0, p_1, p_2 for the cases given in Section 5, Exercises 4-9. For the regular matrices verify that these are the limiting probabilities there obtained. What do p_0, p_1, p_2 mean for the absorbing chains?

7. MARRIAGE RULES IN PRIMITIVE SOCIETIES

In some primitive societies we find rigid rules as to when marriages are permissible. These rules are designed to prevent very close relatives from marrying. The rules can be given precise mathematical formulation in terms of permutation matrices. Our discussion is based, in part, on the work of André Weil and Robert R. Bush.

The marriage rules we find in these societies are characterized by the following axioms.

Axiom 1. Each member of the society is assigned a marriage type.

Axiom 2. Two individuals are permitted to marry only if they are of the same marriage type.

Axiom 3. The type of an individual is determined by the individual's sex and by the type of his parents.

Axiom 4. Two boys (or two girls) whose parents are of different types will themselves be of different types.

Axiom 5. The rule as to whether a man is allowed to marry a female relative of a given kind depends only on the kind of relationship.

Axiom 6. In particular, no man is allowed to marry his sister.

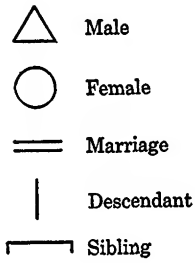
Axiom 7. For any two individuals it is permissible for some of their descendants to intermarry.

Example. Let us suppose that there are three marriage types, t_1, t_2, t_3. Two parents in a given family must be of the same type, since only then are they allowed to marry. Thus there are only three logical possibilities for marriages. For each case we have to state what the type of a son or a daughter will be.

Type of both parents	*Type of their son*	*Type of their daughter*
t_1	t_2	t_3
t_2	t_3	t_1
t_3	t_1	t_2

We must verify that all the axioms are satisfied. Some of the axioms are easy to check (see Exercise 1), others are harder to verify. We will prove a general theorem which will show that this rule satisfies all the axioms.

In order to give a complete treatment to this problem, we must have a simple systematic method of representing relationships. For this we use family trees, as drawn by anthropologists. The following symbols are commonly used:



In Figure 9 we draw four family trees, representing the four kinds of first-cousin relationships between a man and a woman.

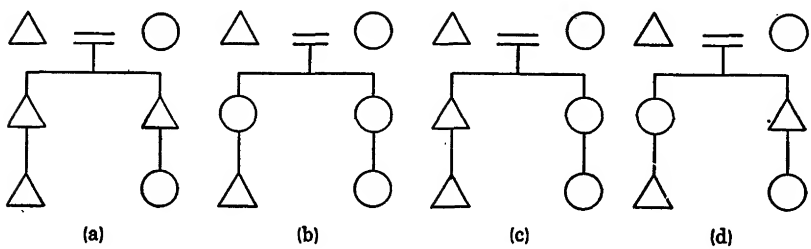


Figure 9

Example (continued). Does our rule allow marriage between a man and his father's brother's daughter? This is the relationship in

Figure 9(a). There are three possible types for the original couple (the grandparents) and in Figure 10 we work out the three cases. We

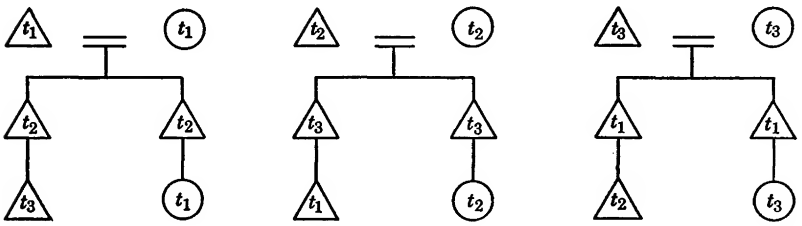


Figure 10

find in each case that the man and woman are of different type, hence such marriages are *never* allowed. Can a man marry his mother's

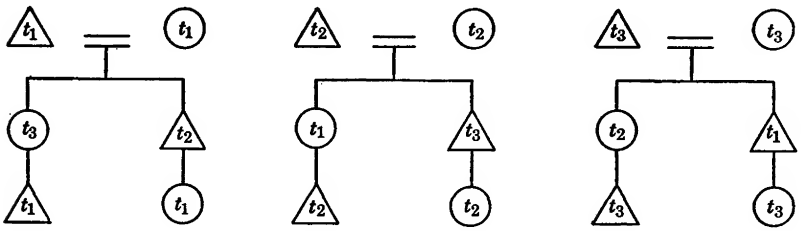


Figure 11

brother's daughter? This is the relationship in Figure 9(d). The three cases for this relationship are found in Figure 11. We find that such marriages are *always* allowed.

We are now ready to give the rules a mathematical formulation. The society chooses a number, say n, of marriage types (Axiom 1). We call these $t_1, t_2, \ldots, t_n$. Our rule has two parts, one concerning sons, one concerning daughters. Let us consider the marriage type of sons. The parents must be of the same marriage type (Axiom 2). We must assign to a boy a type which depends only on the common type of his parents (Axiom 3). If his parents are of type t_i, he will be of type t_j. Furthermore, if some other boy has parents of a type different from t_i, then the boy will be of type different from t_j (Axiom 4). This defines a *permutation* of the marriage types (see Chapter V,

Section 10); the type of a son is obtained from the type of his parents by a permutation specified by the rule of the society. Hence we form the type vector $t = (t_1, \ldots, t_n)$ and represent the permutation in question by the $n \times n$ permutation matrix S. If the type of the parents is component i of t, the type of their sons is component i of tS. By a similar argument we arrive at the permutation matrix D giving the type of daughters.

We have shown that the mathematical form of the first four axioms is to introduce the row vector t and the two permutation matrices S and D. The last three axioms restrict the choice of S and D. This will be considered in the next section.

We have repeatedly seen how the vector and matrix notation allows us to replace a series of equations by a single one. In the present problem this notation allows us to work out a given kind of relationship for *all* marriage types in a single diagram. As a matter of fact, this can be done without knowing how many types there are in the given society, or knowing what the rules are. Let us illustrate this in terms of Figure 11. The couple at the top of the tree is of a given type, represented by our vector t. Their son is of type tS, their daughter of type tD. Then the son of a son is of type tSS, the son's daughter is of type tSD, etc. We arrive at the single vector diagram of Figure 12. If in this figure we take t to have three components, then the diagram is a shorthand for the three diagrams of Figure 11.

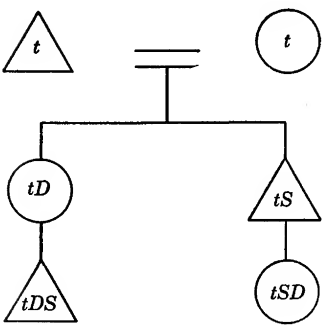


Figure 12

Example (continued). Our t vector is (t_1,t_2,t_3) and

$$D = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}.$$

We know from Figure 11 that a man is always allowed to marry his mother's brother's daughter. Can we see this in Figure 12? The marriage will always be permitted if tDS always equals tSD, which is

equivalent to the matrix equation $DS = SD$. It so happens for our S and D that this equation is correct. But we can see more from Figure 12. No matter how many types there are, this kind of marriage will be permitted if and only if $SD = DS$, i.e., if the two matrices commute.

We have now seen one example of how the nature of S and D determines which kinds of relatives are allowed to marry. This question will be the subject of the next section.

EXERCISES

1. In the example above, verify that the rule satisfies Axioms 1, 3, and 4.

2. In the example above, verify that the matrices S and D given represent the rule given.

3. Construct a diagram for the brother-sister relationship.

4. Using the diagram of Exercise 3, show that, in the above example, brother-sister marriages are never permitted.

5. Find the condition on S and D that would always allow brother-sister marriages. [*Ans.* $S = D$.]

In the *Kariera* society there are four marriage types, assigned according to the following rules:

Parent type	*Son type*	*Daughter type*
t_1	t_3	t_4
t_2	t_4	t_3
t_3	t_1	t_2
t_4	t_2	t_1

Exercises 6-11 refer to this society.

6. Find the t, S, and D of the Kariera society.

7. Show that brother-sister marriages are never allowed in the Kariera society.

8. Show that S and D commute. What does this tell us about first-cousin marriages in the Kariera society?

9. Show that first cousins of the kinds in Figure 9(a) and (b) are never allowed to marry in the Kariera society.

10. Show that first cousins of the kind in Figure 9(c) are always allowed to marry in the Kariera society.

11. Find the group generated by S and D of the Kariera society. (See Chapter V, Section 11.)

In the *Tarau* society there are also four marriage types. A son is of the same type as his parents. A daughter's type is given by:

Parent type	*Daughter type*
t_1	t_4
t_2	t_1
t_3	t_2
t_4	t_3

Exercises 12-17 refer to this society.

12. Find the t, S, and D of the Tarau society.

13. Show that brother-sister marriages are never allowed in the Tarau society.

14. Show that S and D commute. What does this tell us about first-cousin marriages in the Tarau society?

15. Show that first cousins of the kinds in Figure 9(a) and (b) are never allowed to marry in the Tarau society.

16. Show that first cousins of the kind in Figure 9(c) are never allowed to marry in the Tarau society.

17. Find the group generated by S and D of the Tarau society. (See Chapter V, Section 11.)

8. THE CHOICE OF MARRIAGE RULES

In the last section we saw that the marriage rules of a primitive society are determined by the vector t and the matrices S and D. The axioms make no mention of the number of types, and indeed, we will find that we can have any number of types, as long as $n > 1$. But we will find that the choice of S and D are severely limited. This shows that the rules of existing primitive societies required considerable ingenuity for their construction.

We must now consider the last three axioms. For Axiom 5 we need a simple way of describing a kind of relationship. The family tree is our basic tool, but we want to replace the family tree by a suitable matrix.

Let us consider Figure 12. Instead of starting with the grandparents and finding the types of the grandson and the granddaughter, we could start with the grandson, work up to the grandparents, and then down to the granddaughter. For this we must consider how we work "up." If a parent is of type t, the son is of type tS. Hence, if the son is of type t, then the parent is of type tS^{-1} (see Chapter V, Section 10). Similarly, if a daughter has type t, her parents have type tD^{-1}. In Figure 13 we find the new version of Figure 12.

It is easily seen that we can follow this procedure for any relationship. Given a kind of relationship, it determines a matrix M such that if the male of the relationship is of type t, then the female is of type tM. From Figure 13 we see that for "mother's brother's daughter" $M = S^{-1}D^{-1}SD$. We will speak of M as the *matrix of the relationship*. These matrices are all products of S, D, and their inverses, hence each matrix is an element of the group generated by S and D.

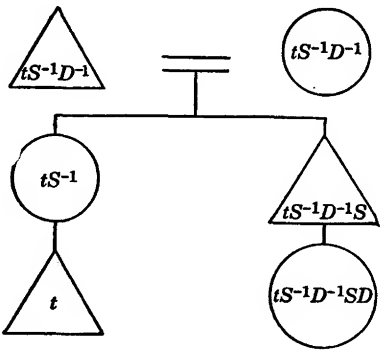


Figure 13

Let us consider Axiom 5. Given any kind of relationship between a man and a woman, we form the matrix of the relationship M. The man will be permitted to marry this relation of his if and only if his type is the same as hers, i.e., if a certain component of t is the same as the corresponding component of tM. This means that this component is left unchanged by the permutation M, which proves our first theorem. (See Chapter V, Section 11.)

Theorem 1. A man is allowed to marry a female relative of a certain kind if and only if his marriage type does not belong to the effective set of the matrix of the relationship.

A second result follows from this theorem easily.

Theorem 2. Marriage between relatives of a given kind is always permitted if the matrix of the relationship has an empty effective set; it is never permitted if the matrix has a universal effective set.

Theorem 3. Axiom 5 requires that in the group generated by S and D every element except I is a complete permutation.

Proof. The axiom states that for a given relationship the marriage must always be allowed or must never be allowed. Hence, by Theorem 2, the matrix of the relationship must have an empty effective set or a universal one. The former means that the matrix is I, the latter that it is a complete permutation (see Chapter V, Section 11). Hence the matrix of every relationship must either be I or a complete permutation matrix. The matrices are elements of the group generated by S and D. And given any element of this group, which can be written as a product of S's and D's, we can draw a family tree having this matrix. Hence the matrices of relationships are all the elements of the group. This means that all the elements of the group, other than the identity, must be complete permutations. This completes the proof.

Theorem 4. Axiom 6 requires that $S^{-1}D$ be a complete permutation.

This theorem is an immediate consequence of the fact that the matrix of the brother-sister relationship is $S^{-1}D$.

Theorem 5. Axiom 7 requires that for every i and j there be a permutation in the group which carries t_i into t_j.

Proof. Let us choose two individuals, one of type t_i and one of type t_j. There must be a descendant of the former who can marry a descendant of the latter. Hence the two descendants must have the same type. This means that we have permutations M_1 and M_2 such that t_i is carried by M_1 into the same type as t_j by M_2. Then $M_1M_2^{-1}$ carries t_i into t_j. Hence the theorem follows.

We have now translated Axioms 5-7 into the following three conditions on S and D: (1) The group generated by S and D consists of I and of complete permutations. (2) $S^{-1}D$ is a complete permutation. (3) For every pair of types there is a permutation in the group that carries one type into the other.

DEFINITION. A permutation group is called *regular* if (a) it is complete, i.e., every element of the group other than I is a complete permutation and if (b) for every pair from among the n objects there is a permutation in the group that carries one into the other.

Basic theorem. To satisfy the axioms we must choose two different $n \times n$ permutation matrices S and D which generate a regular permutation group.

Proof. Conditions (1) and (3) above state precisely that the group generated by S and D be regular. In a regular group every element other than I is a complete permutation; hence condition (2) requires only that $S^{-1}D$ be different from I. Since $S^{-1}D = I$ is equivalent to $D = S$, we need only require that $D \neq S$. This completes the proof.

It is important to be able to recognize regular permutation groups. Here we are helped by a very simple, well-known theorem: A subgroup of the group of permutations of degree n is regular if and only if it has n elements and is complete.

This leads to a relatively simple procedure. We choose n. Then we must pick a group of $n \times n$ permutation matrices which has n elements and is complete, and select two different elements which generate the group. This is always possible if $n > 1$ (see Exercise 11). One of these is chosen as S and one as D. Since there are not very many regular permutation groups for any n, the choice is very limited.

Example. Let us find all possibilities for a society having four marriage types. First of all we must find the regular subgroups of the symmetric group of degree 4, i.e., the groups of permutations on four objects that have four elements and are complete.

Among these we find cyclic groups. Any two of these groups have the same structure and hence lead to equivalent rules. Let us suppose that we choose the permutation group generated by

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}.$$

The group consists of P, P^2, P^3, and I. Either P or P^3 generates the group, and they play analogous roles. We may therefore assume that P is one of the two permutations chosen. This allows us (P,P^2), (P,P^3), and (P,I) as possibilities. We must still ask which is S and which is D. In the second case it makes no difference, since P and P^3 play analogous roles in the group, but there is a difference in the first two cases. This leads to five possibilities:

1. $S = P$, $D = P^2$
2. $S = P^2$, $D = P$
3. $S = P$, $D = P^3$
4. $S = P$, $D = I$
5. $S = I$, $D = P$. This is the Tarau society.

There is only one noncyclic complete subgroup with four elements, consisting of I and the three permutations which interchange two pairs of elements. In this group we have essentially only one case, since all three permutations play the same role.

6. The Kariera society. (See exercises after the last section.)

Two of these six possibilities are actually exemplified in known primitive societies.

EXERCISES

1. Figure 13 shows the matrix of one of the first-cousin relations. Find the matrices of the other three first-cousin relationships.

2. Prove that marriage between relations of a certain kind is permitted if and only if the matrix of the relation is I.

3. Use the result of Exercise 2 to prove that no society allows the marriage between cousins of the types in Figure 9(a) and (b).

4. Which of the six rules described above (in the example) allow marriage between a man and his father's sister's daughter? [*Ans.* 3, 6.]

5. Show that all six rules given in the example above allow marriages between a man and his mother's brother's daughter.

6. There are eight kinds of second-cousin relationships between a man and a woman. Draw their family trees.

7. Find the matrices of the eight second-cousin relationships.

8. Are there any second-cousin relationships for which marriage is forbidden by all possible rules? [*Ans.* Yes.]

9. Test the second-cousin relationships (other than those found in Exercise 8) for each of the six rules given in the example above.

10. For n objects, consider the permutation that carries object number i into position $i + 1$, except that the last object is put into first place. Show that the cyclic group generated by this permutation is regular.

11. Use the result of Exercise 10 to show that a society can have any number of marriage types, as long as the number is greater than one.

12. In the Example of Section 7, prove that S and D generate a regular permutation group.

13. Prove that the following matrices lead to a rule satisfying all axioms.

$$S = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}, \qquad D = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \end{pmatrix}.$$

14. Prove that the rule given in Exercise 13 allows no first-cousin marriages.

9. MODEL OF AN EXPANDING ECONOMY

The following model is a modification of a model proposed by John von Neumann. It is designed to study an economy which is expanding at a fixed rate, but which is otherwise in equilibrium. The model makes certain assumptions about how an economy behaves in equilibrium. These assumptions are idealizations, and it is to be expected that the model will eventually be replaced by a better model. For the present many economists consider the von Neumann model to be a reasonable approximation of reality. Our interest in the model is purely to illustrate how finite mathematics is used in an economic problem.

The economy is described by *n goods* and *m processes*. A good may be steel, coal, houses, shoes, etc. Goods are the materials of production in the economy. Each good may be measured in any convenient units, as long as the units are fixed once and for all. It is convenient to be able to talk of arbitrary multiples of these units; e.g., we will consider not only 2.75 tons of steel but also 2.75 houses. The latter may be interpreted as an average.

A manufacturing process needs certain goods as raw materials (the *inputs*) and produces one or more of our goods (the *outputs*). As a process we may, for example, consider the conversion of steel, wood, glass, etc. into a house. Of course this process may be used to manufacture more than one house, and hence we have the concept of the *intensity* with which a process is used. One of the basic assumptions

is one of linearity, i.e., that k houses will require k times as much of each raw material. Thus we choose an arbitrary "unit intensity" for each process, and the process is completely described if we know the inputs necessary for this unit operation and the outputs produced.

Process number i when operating at unit intensity will require a certain amount of good j as an input. This amount will be called a_{ij}. (In particular, if good j is not needed for process i, then $a_{ij} = 0$.) We will call b_{ij} the amount of good j produced by process i. Here we allow a process to produce several different goods (e.g., a principal output and by-products). But, of course, we allow processes that produce only one good. Then all the b_{ij} for this i will be 0, except for one. The a_{ij} and b_{ij} are nonnegative numbers.

We define the matrix A to be the $m \times n$ matrix having components a_{ij}, and B to be the $m \times n$ matrix with components b_{ij}. Then the entire economy is described by these two matrices.

We must still consider the element of time. It is customary to think of the economy as working in stages or cycles. In one such stage there is just time enough for process i to convert the inputs a_{ij} to outputs b_{ij}. Then in the next stage, these outputs may in turn be used as inputs. The length of this cycle may be any time interval convenient for the study of the particular economy. It may be a month, a year, or a number of years.

Example. Let us take as our economy a chicken farm. Our goods are chickens and eggs, with one chicken and one egg being the natural units. Our two processes consist of laying eggs and hatching them. Let us assume that in a given month a chicken lays an average of 12 eggs if we use it for laying eggs. If used for hatching, it will hatch an average of four eggs per month. From this information we can construct A and B.

Our cycle is of length one month. Good 1 is "chicken," good 2 is "egg," process 1 is "laying," and process 2 is "hatching." The unit of intensity of a process will be what one chicken can do on the average in a month. The input of process 1 is one chicken, i.e., one unit of good 1. The output will consist of a dozen eggs *plus* the original chicken. (We must not forget this, since the original chicken can be used again in the next cycle.) Hence the output is one unit of good 1 and 12 units of good 2. In process 2 the inputs are one chicken

and four eggs, while the output consists of five chickens (the original one plus the four hatched). Hence our matrices are

$$\begin{array}{lc} & \begin{array}{cc}\text{Chicken} & \text{Egg}\end{array} \\ \begin{array}{l}\text{Laying eggs:}\\ \text{Hatching eggs:}\end{array} & A = \begin{pmatrix}1 & 0\\ 1 & 4\end{pmatrix}, \end{array} \qquad \begin{array}{cc} \begin{array}{cc}\text{Chicken} & \text{Egg}\end{array} \\ B = \begin{pmatrix}1 & 12\\ 5 & 0\end{pmatrix}. \end{array}$$

Suppose that our farmer starts with three chickens and eight eggs ready for hatching. He will need two chickens for hatching the eight eggs, and this leaves him one for laying eggs. Hence he uses process 1 with intensity 1, process 2 with intensity 2. We symbolize this by the vector $x = (1,2)$. Note that his inputs are the components of xA. His one laying chicken will lay 12 eggs. He will end up with his original three chickens plus eight new ones. Hence he will have an output of 11 units of good 1 and 12 units of good 2. These are the components of xB. Of his 11 chickens only three can be used for hatching, hence he will employ intensities (8,3). The outputs will be $(8,3)B = (23,96)$, as can easily be checked (see Exercise 1). He now has 96 eggs and only 23 chickens, so that some eggs must go unhatched.

On the other hand, suppose that he starts with only two chickens and four eggs. He will then use intensity (1,1). His laying chicken lays 12 eggs, and with four newly hatched chickens he has a total of six chickens. This result is also given by $(1,1)B = (6,12)$. He now has tripled both his chickens and his eggs. He can use intensity (3,3) on the next cycle, yielding $(3,3)B = (18,36)$, which again triples both the chickens and the eggs. Thus he can continue to use the same proportion of the processes, and will continue to triple his output on every cycle. This economy operates in *equilibrium*.

As was seen in the example, the natural way to represent the intensities of our processes is by means of a row vector. Let x_i be the intensity with which process number i is operated, then the *intensity vector* x is $(x_1, \ldots, x_m)$. Matrix multiplication is then an easy way of finding the total amount of each good needed, and the totals produced. Component j of xA is the sum $x_1a_{1j} + \ldots + x_ma_{mj}$; where x_1a_{1j} is the amount of good j we are using in process 1, x_2a_{2j} the amount we use in process 2, etc. Hence the jth component of xA is the total amount of good j needed in the inputs. Similarly, xB gives the total amounts of the various goods in the outputs.

We must now introduce prices for the various goods. Let y_j be the price of a unit of good j; this must be nonnegative, but it may be zero. (The latter represents a good that is so cheap as to be "practically free.") It is assumed that k units of good j will cost ky_j. The *price vector* y is the column vector $\begin{pmatrix} y_1 \\ \cdot \\ \cdot \\ \cdot \\ y_n \end{pmatrix}$. Let us consider the products Ay and By. In Ay the ith element is $a_{i1}y_1 + \ldots + a_{in}y_n$; the product $a_{i1}y_1$ is the amount of good 1 needed for unit operation of process i multiplied by the per unit price of good 1, hence this is the cost of good 1 used in the process, $a_{i2}y_2$ is the cost of good 2 used, etc. Hence the ith component of Ay is the total cost of inputs for a unit intensity operation of process i. Similarly, By gives the cost (value) of the outputs.

Finally, we consider the products xAy and xBy. Since x is $1 \times m$, the matrices $m \times n$, and y is $n \times 1$, each product is 1×1—or a number. An analysis similar to those above shows that xAy is the total cost of inputs if the economy is operated at intensity x, with prices y, and xBy is the total value of all goods produced. (See Exercise 2.)

Example (continued). Suppose that a chicken costs 10 monetary units, while an egg costs 1 unit; then $y = \begin{pmatrix} 10 \\ 1 \end{pmatrix}$. Here

$$Ay = \begin{pmatrix} 10 \\ 14 \end{pmatrix} \quad \text{and} \quad By = \begin{pmatrix} 22 \\ 50 \end{pmatrix}.$$

This means that process 1, laying eggs, multiplies our investment by a factor of 2.2; while process 2, hatching, brings in over \$3.50 for every \$1.00 invested. There will be pressure to use the hens just for hatching—which will create a shortage of eggs, bringing about a drastic change in prices. Suppose now that a chicken costs only six times as much as an egg, i.e., $y = \begin{pmatrix} 6 \\ 1 \end{pmatrix}$. Then

$$Ay = \begin{pmatrix} 6 \\ 10 \end{pmatrix} \quad \text{and} \quad By = \begin{pmatrix} 18 \\ 30 \end{pmatrix}.$$

In this case each process triples our investment, and there will be no undue monetary pressure. Hence the farmer can set up his processes so as to be in equilibrium, and the price structure will be stable.

The remaining factor to be considered is the expansion of the economy. We assume that everything expands at a constant rate, i.e., that there is a fixed *expansion factor* α such that if the processes operate at intensity x in this cycle, they operate at intensity αx during the next cycle, $\alpha^2 x$ after that, etc. There is also something similar to expansion for the money of the economy, namely, that through bearing interest, y units of money in this cycle will be worth βy units after the cycle. We again assume that the *interest factor* β is fixed once and for all in equilibrium. Usually these factors will be greater than 1, but this does not have to be the case. Thus $\alpha = 1$ represents a stationary economy, and $\alpha < 1$ represents a contracting economy.

This completes the survey of the basic concepts. We must now lay down our assumptions concerning the behavior of an economy which is in equilibrium. These assumptions serve as axioms for the system.

First of all, we must assure that we produce enough of each good in each cycle to furnish the inputs of the next cycle. If in a given cycle the economy functions at intensity x, it will function at αx next time. The outputs this time will be xB, while the inputs next time will be αxA; hence we must require

Axiom 1. $$xB \geq \alpha xA.$$

(When we write a vector inequality, we mean that the inequality holds for every component.) We will of course have to require similar conditions for the future. For example, in the second cycle the outputs are αxB, and the inputs needed for the third cycle are $\alpha^2 xA$. But when we write the condition that the former be greater than the latter, an α cancels, and we have again the same condition as in Axiom 1. Hence this axiom serves for all cycles.

The first condition assures that it is possible for the economy to expand at the constant rate α. We must also assure that the economy is financially in equilibrium. Suppose that the output of some process was worth more than β times the input. Then we would be prepared to pay interest at a larger rate to some one willing to invest in our process. Hence β would increase. Thus, in equilibrium this must not be possible; no process can produce profits at a rate greater than that

given by investment. If we operate processes at a unit intensity, then Ay gives the costs of inputs, while By gives the cost of outputs. The latter cannot exceed the former by more than a factor β for any process

Axiom 2. $$By \leq \beta Ay.$$

The next assumption concerns surplus production. If we produce more of a given good than can be used by the total economy, the price drops sharply as merchants try to get rid of their produce. It is customary to assume, for the sake of simplicity, that such goods are free, i.e., to give them price zero. The vector difference $xB - \alpha xA = x(B - \alpha A)$ gives the amounts of overproduction, i.e., the jth component is positive if and only if good j is overproduced. If we assign price zero to these goods, then in the product of the above vector with y every nonzero factor of the former is multiplied by zero; hence the product of the two vectors will be 0.

Axiom 3. $$x(B - \alpha A)y = 0.$$

Now we turn to the question of whether a given process is worth undertaking. From Axiom 2 we know that no process can yield more profit than investment can. But if it yields any less, it is better not to use it, but rather to invest our money. Hence in Axiom 2 we form the difference $By - \beta Ay$; if the ith component of this is negative, process i should not be used; it must be assigned intensity 0. Similar to the argument used for Axiom 3, this shows that multiplying this vector difference by x must yield zero.

Axiom 4. $$x(B - \beta A)y = 0.$$

Our final assumption is that something worth while is produced in the economy, i.e., that the value of all goods produced is a positive amount.

Axiom 5. $$xBy > 0.$$

If for a given economy (given A and B) we find vectors x and y and numbers α and β which satisfy these five axioms, we say that we have found a *possible equilibrium solution* for the economy.

Example (continued). We have already seen that if $x = (1,1)$, the economy expands at the fixed rate $\alpha = 3$. We can now check that

Axiom 1 is satisfied. Actually, xB turns out to equal αxA. Similarly, we have noted a monetary equilibrium if $y = \begin{pmatrix} 6 \\ 1 \end{pmatrix}$, and each process multiplies the money put into it by a factor of $\beta = 3$. We can check that Axiom 2 holds. Actually By is equal to βAy in this case. From these two equations we also know that $x(B - \alpha A)$ and $(B - \beta A)y$ are identically 0; hence Axioms 3 and 4 hold. Finally, $xBy = 48$; the total value of goods produced is positive, so that Axiom 5 holds. Therefore these values of x, y, α, and β represent an equilibrium for the economy. It can also be shown that these are the only possible values of α and β, and that x and y must be proportional to those shown here (which may be thought of simply as a change in the units).

In our example we found one and only one equilibrium for the economy, and we found that $\alpha = \beta$. This raises several very natural questions: (1) Is there a possible equilibrium for every economy? (2) If yes, then is there only one? (3) Must the expansion factor always be the same as the interest factor? In the next section we will establish the following answers: (1) For every economy satisfying a certain restriction (which is certainly satisfied for all real economies) there is a possible equilibrium. (2) There may be more than one equilibrium, though the number of different possible expansion factors is finite. (In the example there is essentially only one possibility for x and y; however this is not true in general.) (3) The interest and expansion factors are always equal in equilibrium.

EXERCISES

1. In the example, for $x = (1,2)$, verify for three cycles that xA and xB give the correct inputs and outputs.

2. Give an interpretation of xAy and xBy:
(a) Using the interpretations of xA and xB given above.
(b) Using the interpretations of Ay and By given above.
(c) And show that the results in (a) and (b) are the same.

3. In the example suppose that two chickens lay eggs and three hatch eggs. Find x, xA, and xB. Substitute these quantities into Axiom 1, and find the largest possible expansion factor. [*Ans.* $\alpha = 2$.]

4. In the example, suppose that chickens cost 80 cents and eggs cost five

cents. Find y, Ay, and By. Substitute these quantities into Axiom 2, and find the smallest possible interest factor. [*Ans.* $\beta = 4$.]

5. Show that the x, y, α, and β found in two previous Exercises do *not* lead to equilibrium, by showing that Axioms 3 and 4 fail to hold.

6. Show that if $\alpha = \beta = 3$, then the only possible x's and y's are proportional to those given above. (*Hint:* show that the axioms force us to choose $x_1 = x_2$ and $y_1 = 6y_2$.)

The remaining problems refer to the following economy: On a chicken farm there is a breed of chicken that lays an average of 16 eggs a month, and such that they can hatch an average of $3\frac{1}{5} = \frac{16}{5}$ eggs.

7. Set up the matrices A and B.

8. Suppose that three chickens lay and five chickens hatch. Find x, xA, and xB. What is α? [*Ans.* $x = (3,5)$; $xA = (8,16)$; $xB = (24,48)$; $\alpha = 3$.]

9. Suppose that chickens cost 40 cents and eggs five cents. Find y, Ay, and By. What is β?

10. Verify that the x, y, α, and β found in the previous exercises represent an equilibrium for the economy, by substituting these into the five axioms.

11. Suppose that we start with 16 chickens and 32 eggs. Choose the intensities so that the economy will be in equilibrium, and find what happens in the first three months. [*Ans.* $x = (6,10)$; 432 chickens, and 864 eggs.]

12. Suppose that with 16 chickens and 32 eggs (see Exercise 11) we start out by having only five hatching, the others laying. Show that we cannot have as many chickens after three months as we would have in the equilibrium solution.

10. EXISTENCE OF AN ECONOMIC EQUILIBRIUM

We must ask whether the axioms can always be satisfied, i.e., whether the model of the economy allows such an equilibrium.

Of course we are interested only in an economy that could really occur. That means that these goods must be goods that are somehow produced, and that they cannot be produced out of nothing. Hence every process must require at least one raw material and every good has at least one process that produces it. We summarize this:

Restriction. Every row of A and every column of B has at least one positive component.

Theorem. If A and B satisfy the restriction, then an equilibrium is possible.

We will sketch the proof of this theorem. From Axiom 3 we have that $xBy = \alpha xAy$, while from Axiom 4, $xBy = \beta xAy$. Hence $\alpha xAy = \beta xAy$. Furthermore, from Axiom 5 we know that xBy is not zero, hence xAy is not zero. Then $\alpha = \beta$. Hence *in equilibrium the rate of expansion equals the interest rate.*

If $\alpha = \beta$, then Axioms 3 and 4 are equivalent. We can also rewrite the first two axioms (using our result):

Axiom 1'. $\quad x(B - \alpha A) \geq 0.$

Axiom 2'. $\quad (B - \alpha A)y \leq 0.$

If we multiply the first inequality by y on the right, and the second by x on the left, we see that Axiom 3 (and hence 4) follows from these two axioms. Hence we need only worry about Axioms 1', 2', and 5.

The key to the proof is to reinterpret the problem as a game-theoretic one. This is done in spite of the fact that no game is involved in the model. We simply use the mathematical results of the theory of games as tools.

Axioms 1' and 2' suggest that we think of the matrix $B - \alpha A$ as a matrix game. We would then like to think of the vectors x and y as mixed strategies for the two players. The vectors are nonnegative, but the sum of their components need not be 1. However, we know that multiplying x by a constant can be thought of as a change in the units of intensities, and multiplying y by a constant is equivalent to a change in the units of the various goods. Hence, without loss of generality, we may assume that x and y have component sum 1, and think of them as mixed strategies. If we do this, the two axioms state precisely that the game has value zero, and that x and y form a pair of optimal strategies for the two players. Thus our first problem is to choose α so that the "game" $B - \alpha A$ has value zero.

Example 1. Let us set up the example of the last section as a game.

$$M = B - \alpha A = \begin{pmatrix} 1 - \alpha & 12 \\ 5 - \alpha & -4\alpha \end{pmatrix}.$$

If we choose $x = (\frac{1}{2}, \frac{1}{2})$ as a mixed strategy for the row player, then $xM = [3 - \alpha, 2(3 - \alpha)]$. If $\alpha < 3$, the components are both posi-

tive; hence the game has value greater than zero. If we choose $y = \begin{pmatrix} 6 \\ 1 \end{pmatrix}$ as a mixed strategy for the column player, then

$$My = \begin{bmatrix} 6(3 - \alpha) \\ 10(3 - \alpha) \end{bmatrix}.$$

If $\alpha > 3$, both components are negative, and hence the game has negative value. We thus see that the only value of α that could possibly give us a zero value of the game is $\alpha = 3$, and we see from the above that in this case the value really is zero, and x and y are optimal strategies. (See Exercise 1.)

We must now show that the above example is typical in that we can always find an α making the value of $B - \alpha A$ equal to zero. We may write this matrix as the sum $B + \alpha(-A)$, and think of our game as a combination of game B and game $(-A)$.

By our restriction, every column of B has a positive entry. The strategy vector y for the column player must have at least one positive component. Hence in the product By, one of the components at least must be positive. Hence the value of the game B is positive. Since every row of A has a positive entry, every row of the game $-A$ must have a negative entry. Hence at least one component of $x(-A)$ must be negative, and hence $(-A)$ has a negative value.

In the combination $B + \alpha(-A)$ the second term is negligible for very small α; hence for these the game has positive value. As α increases, we keep adding larger negative quantities to some of the entries of the game, i.e., we keep decreasing some of these entries. Hence the value of the game decreases steadily. For very large α the first term is negligible, and hence the combined game has negative value. For some intermediate value of α the game must have value zero.

Example 1 (continued). The value of the combined game M is plotted for various α in Figure 14. Since B has value $\frac{15}{4}$ and $-A$ has value -1 (see Exercise 2), at the beginning the game M has value nearly $\frac{15}{4}$, and near the end it has value nearly $2 - \alpha$, which is well below zero (see Exercise 3).

We know that there is at least one α for which the game $B - \alpha A$ has value zero. By choosing such an α together with a pair x, y of

optimal strategies, we arrive at a set of quantities satisfying Axioms 1′ and 2′. This still leaves the question of Axiom 5.

If there are two values of α, say $p < q$, for which the game has value zero, every value between p and q also has this property. This

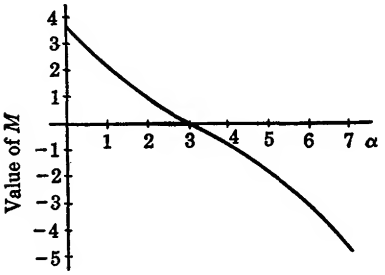


Figure 14

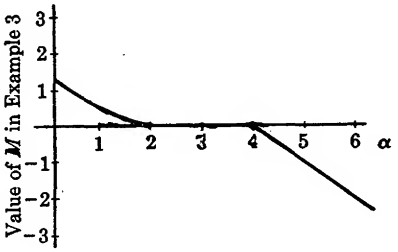


Figure 15

is because the value of the game cannot increase as α increases, as we saw above. Hence we must have a situation such as that shown in Figure 15. It can be shown, however, that most of these values represent methods of procedure where nothing worth while is produced, i.e., where Axiom 5 fails. For Axiom 5 to hold, different values of α can be achieved only by using at least one new process. Since there are only a finite number of processes, we can have only a finite number of different possible α's on the interval between p and q. If p is the smallest possible expansion rate and q the largest, then p and q are such that Axiom 5 can be satisfied, and there may be a limited number of additional ones in between.

Example 2. In the chemical industry we are interested in manufacturing compounds P, Q, and R. We assume that the basic chemicals are available in plentiful supply, and that their cost can be neglected for this analysis. But to manufacture compound P we must have a unit of both P and Q available, while to manufacture Q we must have P and R available. Compound R is a by-product of both manufacturing processes. The exact quantities are given by

$$\begin{array}{l} \\ \text{Manufacture of } P: \\ \text{Manufacture of } Q: \end{array} \quad \begin{array}{c} P \quad Q \quad R \\ A = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix} \end{array} \qquad \begin{array}{c} P \quad Q \quad R \\ B = \begin{pmatrix} 6 & 0 & 1 \\ 0 & 3 & 2 \end{pmatrix}. \end{array}$$

Then

$$M = B - \alpha A = \begin{pmatrix} 6-\alpha & -\alpha & 1 \\ -\alpha & 3 & 2-\alpha \end{pmatrix}.$$

Let us choose

$$x = (\tfrac{1}{2},\tfrac{1}{2}) \quad \text{and} \quad y = \begin{pmatrix} \frac{1}{6} \\ \frac{1}{3} \\ \frac{1}{2} \end{pmatrix}.$$

Then

$$xM = [3-\alpha,\ \tfrac{1}{2}(3-\alpha),\ \tfrac{1}{2}(3-\alpha)] \quad \text{and} \quad My = \begin{bmatrix} \frac{1}{2}(3-\alpha) \\ \frac{2}{3}(3-\alpha) \end{bmatrix}.$$

From this we see that if $\alpha < 3$, then the row player has a guaranteed profit, while if $\alpha > 3$, the column player does. Thus $\alpha = 3$ is the only possibility, and for this case the value of the game is zero, and the vectors x and y are optimal strategies, as can be seen from the fact that xM and My have all components zero. Thus there is a unique equilibrium, with $\alpha = \beta = 3$.

We also find that the mixed strategy x is unique, which means that the two processes must be used with the same intensity. However, the strategy y is not unique. We may instead use

$$y' = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} \quad \text{or} \quad y'' = \begin{pmatrix} 0 \\ \frac{1}{4} \\ \frac{3}{4} \end{pmatrix}$$

or any mixture $ty' + (1-t)y''$, $0 \leq t \leq 1$. Our y is the case $t = \frac{1}{3}$. Hence we see that different price structures are possible, each leading to the same expansion rate.

Example 3. This "economy" is a schematic representation of the production of essentials and inessentials in a society. Goods are lumped together into two types, E (essential goods) and I (inessential goods or luxury items). For the manufacture of E we need only essential goods (since anything so needed is essential). For the manufacture of I we may need both types of raw materials. Let us suppose that our economy functions as follows.

	E	I		E	I
Manufacture of essentials:	1	0		4	0
Manufacture of luxuries:	1	1		0	2

$$A = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}, \qquad B = \begin{pmatrix} 4 & 0 \\ 0 & 2 \end{pmatrix}$$

Then

$$M = B - \alpha A = \begin{pmatrix} 4-\alpha & 0 \\ -\alpha & 2-\alpha \end{pmatrix}.$$

With a little patience we can determine the values of M for various values of α, and we arrive at the curve in Figure 15. (See Exercise 4.) Hence α must be between 2 and 4. For $\alpha = 4$, we have the optimal strategies $x = (1,0)$ and $y = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, which satisfy all our axioms; while for $\alpha = 2$ we have

$$x = (\tfrac{1}{2},\tfrac{1}{2}) \quad \text{and} \quad y = \begin{pmatrix} 0 \\ 1 \end{pmatrix}.$$

For in-between values of α we cannot satisfy Axiom 5. (See Exercises 5-7.) Hence there are two possible equilibria: (1) The society can decide to manufacture only essentials, in which case the production of these will increase rapidly. (2) By putting a high enough value on inessentials, it will arrive at an equilibrium in which both essentials and inessentials are produced, but then the rate of expansion is considerably decreased.

We have now provided complete answers for the three questions raised at the end of the last section, providing a mathematical solution to a series of economic problems.

EXERCISES

1. In Example 1 verify that for $\alpha = 3$ the game M has value 0, and that the x and y given are optimal strategies.

2. In Example 1 solve the 2×2 games B and $-A$, finding their values and pairs of optimal strategies.

3. In Example 1:

(a) Show that the game M is nonstrictly determined for every α.

(b) Find the value of M for any α. [*Ans.* $(5 + \alpha)(3 - \alpha)/(4 + \alpha)$.]

(c) Show that the value for $\alpha = .01$ is very near $\frac{15}{4}$.

(d) Show that the value for $\alpha = 100$ is very near -98.

(e) Show that the value is 0 if and only if $\alpha = 3$.

4. Find the value of M in Example 3 for $\alpha = 0, 1, 2, 3, 4, 5$, and 6. (*Hint:* Some of these games are strictly determined.)

[*Ans.* 1.33; .60; 0; 0; 0; −1.00; −2.00.]

5. In Example 3, for $\alpha = 4$, verify that the strategies given are optimal, and that Axiom 5 is satisfied.

6. In Example 3, for $\alpha = 2$, verify that the strategies given are optimal, and that Axiom 5 is satisfied.

7. In Example 3, for $\alpha = 3$, find the unique optimal x and y, and show that Axiom 5 is *not* satisfied. Prove that the same happens for every α if $2 < \alpha < 4$.

The remaining problems refer to the following economy: There are four goods and five processes, and the economy is given by

$$A = \begin{pmatrix} 0 & 0 & 1 & 1 \\ 0 & 0 & 2 & 2 \\ 0 & 4 & 0 & 2 \\ 2 & 1 & 1 & 0 \\ 0 & 1 & 0 & 2 \end{pmatrix}, \qquad B = \begin{pmatrix} 0 & 0 & 4 & 2 \\ 0 & 0 & 5 & 7 \\ 6 & 5 & 4 & 0 \\ 0 & 4 & 0 & 3 \\ 3 & 0 & 6 & 0 \end{pmatrix}.$$

Also let $x = (\frac{1}{2},\frac{1}{2},0,0,0)$; $x' = (0,0,\frac{2}{5},\frac{3}{5},0)$;

$$y = \begin{pmatrix} \frac{1}{3} \\ \frac{2}{3} \\ 0 \\ 0 \end{pmatrix}, \qquad y' = \begin{pmatrix} 0 \\ 0 \\ \frac{1}{2} \\ \frac{1}{2} \end{pmatrix}.$$

8. Verify that A and B satisfy the restriction.

9. Compute $M = B - \alpha A$.

10. Compute xM, $x'M$, My, and My'.

11. When will $x'M$ have all positive entries? When will My' have all negative entries? What possibilities does this leave for α?

[*Ans.* $\alpha < 2$; $\alpha > 3$; $2 \leq \alpha \leq 3$.]

12. Show that for the remaining possible values of α the game M has value zero, and x and y are optimal strategies.

13. Show that for the largest possible α the vectors x and y' provide optimal strategies which satisfy Axiom 5.

14. Show that for the smallest possible α the vectors x' and y provide optimal strategies which satisfy Axiom 5.

15. If α is in between its two extreme values, show that:

(a) xM is positive in its last two components, and hence the second player can use only his first two strategies.

(b) My is negative in its last three components, and hence the first player can use only his first two strategies.

(c) For these cases it is impossible to satisfy Axiom 5.

16. Process number five is in a special position. Why? [*Ans.* Never used.]

17. Use the results of Exercises 8–16 to show that there are exactly two possible equilibriums for this economy. Interpret each equilibrium, and point out the differences between the two methods of operating the economy.

[*Ans.* At the price of reducing the expansion rate, the economy can produce a larger variety of goods. To achieve this, the additional types of goods must be valued (relatively) very high.]

SUGGESTED READING

Luce, R. Duncan and Albert D. Perry, "A Method of Matrix Analysis of Group Structure," *Psychometrica*, **14** (1949), pp. 95-116.

Neyman, J., *First Course in Probability and Statistics*, Holt, New York, 1950, Chap. III.

Estes, W. K., and C. J. Burke, "Application of a Statistical Model to Simple Discrimination Learning in Human Subjects," *J. Exptl. Psychol.*, **50** (1955), pp. 81-88.

Bush, R. R., and F. Mosteller, *Stochastic Models for Learning*, Wiley, New York, 1955.

Weil, André, "Sur L'Étude de Certains Types de Lois de Marriage (Système Murngin)," in, Appendice a la Première Partie, *Les Structures Elementaire de la Parente*, by Claude Levi-Strauss, Presses Universitaires de France, Paris, 1949, pp. 278-285.

Kemeny, John G., Oskar Morgenstern, and Gerald L. Thompson, "A Generalization of the von Neumann Model of an Expanding Economy," *Econometrica*, **24** (1956), pp. 115-135.

INDEX

ABSORBING Markov chain, 326
Alternatives (complete set of), 36
Arguments, 42
Aristotle, 1
Associative law for multiplication of matrices, 200
Auxiliary symbols, 2

BASIC conjunction, 16
Basic series circuit, 51
Bayes theorem, 136
Beer and ale experiment, 151
Biconditional, 11
Binary number system, 70, 72
Binomial theorem, 106
Birthday problem, 125
Blocking coalition, 75
Board of Estimate of New York, 77
Boole, George, 1, 69
Boolean algebra, 69
Bridge, 103, 124
Bush, Robert R., 343

CANCER study, 89
Cell, 79
Closed half plane, 250
Chicken farm, 264, 354, 360
Circuit:
  basic series, 51
  switching, 49
Coalitions, 74, 108
  blocking, 75
  losing, 75
  maximal losing, 77
  minimal winning, 75, 110, 111
  voting, 74
  winning, 75, 109
Coin tossing, 116, 121, 131, 144, 148, 150, 158, 159, 161, 166, 170
Column vector, 178
Communication:
  network, 315
  matrices, 315
Commutative group, 240
Commutative law of addition for vectors, 180
Complement (of a set), 59
Complete permutation, 126, 350
Compound statements, 1
  relation to sets, 63
Conclusion (of an argument), 42
Conditional, 10
Conditional probability, 129
Conjunction, 5
  basic, 16
Connectives, 2
Constants, 2
Contradictories, 35
Contrapositive, 39
Contraries, 34
Converse, 39
Convex set, 259
  polygonal, 251
Convex polygons, 254
Craps, 114, 164, 168, 170
Cross-partition, 80
Cyclic group, 244

DEGREE of a permutation matrix, 237
Derived games, 295
Determinant, 214
Dictator, 76, 111
Difference (of sets), 60
Diffusion, 174, 227
Directed graphs, 308
Disjoint subsets, 60
Disjunction, 5
  exclusive, 6
  inclusive, 6
Dominance relation, 307
  matrix of, 308
  one-stage, 309
  two-stage, 309
Doughnuts, dollars to, 120

EFFECTIVE set of a permutation matrix, 246
Empty set, 55

Equation, linear, 205
  locus of, 249
Equilibrium:
  in genetics, 324, 326
  of an economy, 353, 355
Equiprobable measure, 121
Equivalence, 32
Estes learning model, 334
Estes, W. K., 334, 340
Exclusive disjunction, 6
Expansion factor, 357
Expected value, 159, 166
  of a game, 276, 282

FACTOR:
  expansion, 357
  interest, 357
Fallacy, 44
Fixed point, 219
Fundamental theorem of matrix games, 291

GAME matrix, 280
Game theory, 266
Games:
  derived, 295
  expected value of, 276, 282
  fair, 275
  fundamental theorem of, 291
  matrix, 267, 279, 293
  nonstrictly determined, 271
  strictly determined, 266, 283
  symmetric, 289
  value of, 281
Generators of a group, 245
Genes, dominant and recessive, 321
Genetics, stochastic process in, 321
Goods, 353
Group, 240
  commutative, 240
  generators of, 245
  subgroup of, 243
  symmetric, 241

HALF PLANE:
  closed, 250
  open, 250
Hat check problem, 126

IMPLICATION, 30
Inclusive disjunction, 6
Inconsistent statements, 34, 36
  probability of, 117
Independent trials with more than two outcomes, 162
  with two outcomes, 146
Indirect method (of proof), 47
Inequalities, system of, 251
Intensity vector, 355
Interest factor, 357
Intersection (of sets), 58
  laws governing, 69
Inverse of a square matrix, 212

KARIERA society, 347

LAND OF OZ, 231, 235
Law of large numbers, 156
Limiting probabilities:
  Estes model, 339
  for Markov chain, 221
  genetics, 322
Linear equations, 205 ff
  locus of, 249
Linear functions, 233 ff
Linear programming, 249
Linear programming problems, 259
Linear transformations, 219, 234
Locus of a linear equation, 249
  of a linear inequality, 250
Logical possibilities, 19
Logical relations, 30
  onefold, 34
  twofold, 34
Logically false statement, 19
  probability of, 117
Logically true statement, 19
Losing coalition, 75
  maximal, 77

MAJORIZATION (in a game), 294
Markov chains, 147, 171, 194, 217, 224
  absorbing, 326
Marriage types, 343
Matrices, 190
  addition of, 198
  applications to Markov chains, 217
  communication, 315
  dominance, 308
  game, 267, 279, 293
  inverse, 212
  multiplication by constants, 198
  multiplication by vectors, 193
  multiplication of, 199
  of a marriage relationship, 349
  permutation, 236
  regular, 220
  sociometric, 307
  square, 191
  stochastic, 217
  transpose, 240
Measure:
  as an area, 133
  equiprobable, 121
  probability, 114, 116
  tree, 141
Minimal winning coalition, 75, 110, 111
Minorization (in a game), 294

Mixed strategy, 274
Monte Carlo, 167, 171
Morgenstern, O., 302
Morra, 278, 300
Multinomial theorem, 107

NEGATION, 6
Network, switching, 49
von Neumann, John, 302, 353
Nonstrictly determined games, 271
Number of elements in a set, 88
Number systems:
  binary, 70, 72
  ternary, 74

ODDS, 118
Open half plane, 250
Optimal strategy, 268

PARTITIONS, 79
  counting, 96
  ordered, 96
Pascal triangle, 101
Pecking order, 307
Permutation, 92
  complete, 126 (350)
Permutation matrix, 237 ff
  complete, 246
  degree of, 237
  effective set of, 246
  transpose of, 240
Pivot, 109
Poker:
  bluffing in, 304
  hands, 124, 139
  simplified, 302, 305
Polygon, convex, 254
Polygonal convex set, 251
Possibilities, logical, 19
Power, in a dominance situation, 311
Powerless member, 76, 111
Premise, 42
Presidents' birthdays, 128
Price vector, 356
Primary elections, 54
Probability, 113
  conditional, 129
  measure, 114
  of a statement, 114
  relation to odds, 118
  transition, 171
Probability measure, 114
  as area, 133
  equiprobable, 121
  properties of, 116
Processes, economic, 353
Proper subgroup, 243
Proper subset, 55
Pure strategy, 280

RECTANGULAR coordinate system, 186
Reductio ad absurdum, 48
Regular matrix, 220
Regular permutation group, 350
Relations:
  logical, 30
  onefold, 34
  twofold, 34
Roulette, 167
Row vector, 178

SELF-CONTRADICTION, 19
Set, 54
  convex, 259
  cross-partition of, 80
  description of, 54
  empty, 55
  listing of, 54
  number of elements in, 88
  partition of, 79
  polygonal convex, 251
  probability measure of, 114
  relation to compound statements, 63
  unit, 55
  universal, 55
Shapley, L. S., 109
Shubik, M., 109
Simple statement, 1
Simplified poker, 302, 305
Small group behavior, 85
Sociometric matrix, 307
Square matrix, 191
  determinant of a 2 x 2, 214
  inverse of, 212
  powers of, 203
  transpose of, 240
Standard deviation, 159
State, 171
  absorbing, 326
  average number of times in, 327
  average time to return to, 323
Statements:
  compound, 1
  inconsistent, 34, 36
  logically false, 19
  logically true, 19
  probability of, 114
  related, 34
  simple, 1
  truth set of, 64
  unrelated, 34
Stationary Markov process, 220
Stimulus elements, 334
Stochastic process, 139
  in genetics, 321
  matrix, 217
Stone, paper, scissors, 289
Strategy, 268
  mixed, 274, 281

optimal, 268, 281
pure, 280
Strictly determined games, 266, 283
Subcontraries, 34
Subgroup, 243
cyclic, 244
proper, 243
Subsets (of a set), 55
complement of, 59
difference of two, 60
disjoint, 60
intersection of two, 58
laws governing, 69
operations on, 58
proper, 55
union of two, 59
Symmetric game, 289
Symmetric group, 241
System of inequalities, 251
Switching circuits, 49

TARAU society, 348
Ternary number system, 74
T-maze, 84
Transformations:
fixed point of a linear, 219
linear, 219, 234
quadratic, 324
Transition diagram, 172
Transition probability, 171
Transpose of a square matrix, 240
Tree diagram, 26
Tree measure, 141
Truth set (of a statement), 64
Truth tables, 5
Tucker, A. W., 302
Twenty questions, 81

UNION of sets, 59
laws governing, 69
Unit set, 55
United Nations, 76, 110
Universal set, 55

VALID (argument), 42
Value of a game, 281
Variables, 2
Vector(s):
addition of, 179
column, 178
combined with matrices, 190
components of, 178, 179
equality of, 179
multiplication by a constant, 180
multiplication of, 183, 185
probability, 218
row, 178
subtraction of, 181
Venn diagrams, 59
Veto power, 76, 111
Voting coalitions, 74
Voting power, 108

WEIL, André, 343
Winning coalition, 75, 109
minimal, 75, 110, 111
World Series, 27, 79, 146, 166, 170

AL

KEMENY
SNELL
THOMPSON

INTRODUCTION TO FINITE MATHEMATICS

AP
1K47

PRENTICE
HALL